THE UNIVERSITY *of* EDINBURGH

Zoonotic Barriers: Defining Species-specific Porcine Restriction Factors against Coronaviruses

Long Fung (Ocean) Chau

THE UNIVERSITY *of* EDINBURGH

PhD in Infection and Immunity

2024

Declaration

I declare that the thesis has been composed by myself and not others. The work presented in this thesis is my own except for some data in Chapter 5, which results from a collaboration with the Griffiths group at Moredun Research Institute, Edinburgh, UK. Handling of animals including breeding, euthanasia and clinical assessment was carried out by veterinarians and trained professionals based in Moredun and the Large Animal Research and Imaging Facility (LARIF). Histology staining was performed by the Easter Bush pathology department and lung sample histopathology was scored independently by two pathologists, Prof Tanja Opriessnig and Dr David Doward. All other data were generated by myself. The content in this thesis has not previously been submitted for a degree at this or other universities.

Long Fung (Ocean) Chau

Acknowledgements

There is no one I would like to thank, I do this all by myself, I am the hero and I am the best. Well, maybe I should also thank my parents, who bought me the flight ticket from Hong Kong to Edinburgh so I could start my PhD before I received my first stipend.

All jokes aside, I would like to thank my parents for being the best parents ever. For supporting every (lawful) decision I made and never questioning my ability. My supervisor Dr Finn Grey, for giving me all possible support. Saying you are a life mentor is an overstatement, but I have no doubt learning a lot from you outside of science. My second supervisor Dr Christine Tait-Burkard, for your expert knowledge of coronavirus when I am the only person working on it in my own lab group. Thank you for telling me all those wonderful fun facts about Nidovirus.

To everyone in my lab group, it's an absolute pleasure to work and drink with each and every one of you. Matin and Vayalena, thanks for walking through this challenging yet exciting 4-year journey with me. I cannot imagine my PhD life without you two. To Dr. Alex Brown, I forgive you for abandoning me after my first year. Seriously, thank you for guiding me when I have no idea what I am doing. To Dr. Abraham Lee, although

you have chosen the evil path (by joining the industry), your amazing bookkeeping skills have saved us from a lot of trouble. To Dr. Oliver Lin, you are a legend when it comes to lab techniques, without you I would probably need to spend another 4 years to achieve what I have done now. To Dr. Rosemary Blake, the adult in our group when everyone else is acting like a 10-year-old. Thank you for organising all the lab socials and keeping the lab running smoothly.

I would also like to thank everyone from the Digard and Gaunt group, for creating a joyful working environment and initiating wonderful discussions when I am a bit depressed from failing experiments. I cannot possibly hope for better colleagues to share the workplace with.

Last but not least, my wonderful partner Nelly who might be slightly smarter than me. I enjoyed every nerdy discussion and debate we had regarding our projects. You, along with caramel and popcorn, can always pacify me when I have a terrible day with experiments.

Finally, thank you reader, and welcome to the very exclusive club of people-who-have-read-this-thesis.

Table of Contents

Lay summary

Coronaviruses (CoVs) can infect a broad range of animals, from domestic animals like dogs and cats to marine mammals such as seals and beluga whales. They have the ability to 'jump' from one species to another resulting in cross-species transmission events known as zoonosis. Most zoonotic events are 'dead ends', with infection occurring in limited individuals with no subsequent transmissions. However, occasionally, they may result in large-scale outbreaks, as demonstrated by the recent COVID-19 pandemic, caused by SARS-CoV-2, and past epidemics, caused by SARS-CoV-1 and MERS-CoV. This is because CoVs can acquire mutations, allowing them to overcome barriers caused by evolutionary differences between species. One such barrier is the interferon (IFN) response, a critical immune component in fighting off viral infection. Importantly, the IFN response is highly divergent between species.

IFN triggers the simultaneous expression of hundreds of IFN-stimulated genes (ISGs), which are the actual effectors of antiviral immunity. By artificially expressing these effectors, we can pinpoint the function of individual ISGs against different viruses. Previous studies have successfully identified human ISGs that act as barriers against

zoonotic events. However, similar studies in other species are lacking. Pigs are important livestock species that play a critical role in public health, as they host viruses with high pandemic potential, including CoVs. We have therefore developed an ISG expression library for pigs to identify species-specific antiviral effectors for CoVs. In this study, we have applied the library to porcine respiratory coronavirus (PRCV), a CoV that naturally infects pigs, and SARS-CoV-2, which can replicate efficiently in some porcine cell lines but does not infect pigs. We have identified and validated two novel porcine-specific ISGs that potently restrict PRCV and SARS-CoV-2 replication. We have also identified an ISG that restricts SARS-CoV-2 better than PRCV, potentially contributing to barriers against *in vivo* infection. Understanding these porcine barriers may help prevent future disease outbreaks in pigs.

In a separate but related project, we have developed a human ACE2 transgenic pig model that is highly susceptible to SARS-CoV-2 and faithfully replicates COVID-19 in humans. This large animal model shares more similarities to humans than existing models and could serve as a unique tool for developing improved vaccines and therapeutics against COVID-19.

Abstract

The pandemic nature of coronaviruses (CoVs) is underpinned by their ability to mutate and overcome barriers caused by evolutionary differences between species and establish disease in new hosts. One such barrier is the interferon (IFN) response. Systematic analysis using arrayed human IFN-stimulated gene (ISG) expression libraries has been effective at identifying antiviral factors. By comparing antiviral activity against viruses adapted to different species, species-specific restriction factors have been identified that act as barriers against zoonosis. However, most of these studies were focused on humans and libraries for other species are lacking.

Pigs are important livestock species for public health as they are hosts of influenza virus and at least six CoVs. We have generated an arrayed lentiviral-based porcine Type-I ISG expression library comprising 432 ISGs to enable cross-species screening. We hypothesized that a pig endemic virus would be more adept at evading the IFN response of the natural host, compared to a virus, such as SARS-CoV-2, that has not adapted to pigs. Indeed, our results indicate that porcine respiratory coronavirus (PRCV), a highly adapted endemic CoV in pigs that causes pneumonia resembling SARS in humans, was

less sensitive than SARS-CoV-2 to porcine IFN. This suggests species-specific restriction factors have greater antiviral effects against non-adapted viruses. To identify such factors, we have applied the library to PRCV and SARS-CoV-2. We have developed and optimised resources to carry out such screening experiments in porcine cell lines, as existing protocols for human cell lines were not applicable. Using an RT-qPCR-based read-out, we have identified and validated two novel porcine ISGs, ZCWPW1 and PHACTR1, which potently restrict PRCV and SARS-CoV-2 replication. We also showed that the ISG, USF1, restricts SARS-CoV-2 to a greater extent than PRCV.

In addition to the IFN response, receptor compatibility could be a major determinant in the lack of susceptibility of pigs to SARS-CoV-2 *in vivo*. Human ACE2 is the entry receptor for SARS-CoV-2. We have developed a human ACE2 transgenic pig model highly susceptible to SARS-CoV-2, displaying clinical signs, disease progression, and lung inflammation that faithfully replicates COVID-19 in humans. This large animal model could serve as a unique tool in addition to non-human primates and rodents for testing mechanisms of disease and developing improved vaccines and therapeutics against COVID-19.

Abbreviations

2'-O-MTase	2'-O-methyltransferase
ACE2	Angiotensin-Converting Enzyme 2
ADAR1	Adenosine deaminase acting on RNA 1
APN	Aminopeptidase N
ATP	Adenosine triphosphate
CCoV	Canine coronavirus
cDNA	Complementary DNA
cGAS	Cyclic GMP–AMP synthetase
CH25H	Cholesterol-25-hydroxylase
COVID-19	Coronavirus-induced disease 2019
CPE	Cytopathic effect
CRISPR	Clustered regularly interspaced short palindromic repeats
DAD	Diffuse alveolar damage
DMEM	Dulbecco's modified Eagle's medium
DMVs	Double-Membrane Vesicles
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DPI	Day post infection
DPP4	Dipeptidyl Peptidase 4
E protein	Envelope protein
ER	Endoplasmic Reticulum

ERGIC	ER-To-Golgi Intermediate Compartment
FBS	Foetal bovine serum
FCoV	Feline coronavirus
FITC	Fluorescein isothiocyanate
GAF	IFN-gamma activation factor
GAS	Gamma activated sequence
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GOI	Gene of interest
gRNA	Guide RNA
GTase	Guanylyltransferase
GTPase	Guanosine triphosphatases
HPI	Hour post infection
HA	Hemagglutinin
hACE2	Human Angiotensin-Converting Enzyme 2
HBV	Hepatitis B virus
HCMV	Human cytomegalovirus
HCV	Hepatitis C virus
HCoV	Human Coronavirus
HE	Haemagglutinin Esterase
HIV	Human immunodeficiency virus
IBV	Infectious Bronchitis Virus
IC50	Half-maximal inhibitory concentration
ICTV	International Committee on Taxonomy of Viruses
IFI	Interferon gamma inducible protein
IFIT	Interferon-induced protein with tetratricopeptide repeats
IFITM	Interferon-induced transmembrane protein
IFN	Interferon
IFNAR	Interferon α receptors
IFNLR	Interferon λ receptors
IKKε	IκB kinase-ε
IL10R2	Interleukin-10 receptor 2

IRES	Internal ribosome entry site
IRF	Interferon regulatory factor
IRG	Interferon-repressed gene
ISG	Interferon-stimulated gene
ISGF3	Interferon-stimulated growth factor 3
ISRE	Interferon-stimulated response element
JAK1	Janus kinase 1
LY6E	Lymphocyte antigen 6 complex, locus E
M protein	Membrane protein
MAPK	Mitogen-activated protein kinase
MAVS	Mitochondria antiviral signalling protein
MERS-CoV	Middle East Respiratory Syndrome Coronavirus
MDA5	Melanoma differentiation-associated gene 5
MHV	Mouse hepatitis virus
MOI	Multiplicity of infection
MxA	Mycovirus resistance protein A
MYD88	Myeloid differentiation primary response 88
N7-MTase	N7-methyltransferase
N protein	Nucleocapsid protein
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NPTr cell	New-born pig trachea cell
NSK cell	New-born swine kidney cell
Nsp	Non-structural protein
OAS1	2'-5' oligoadenylate synthase 1
ORF	Open reading frame
pACE2	Porcine Angiotensin-Converting Enzyme 2
PAMP	Pathogen-associated molecular pattern
pAPN	Porcine aminopeptidase N
PBS	Phosphate-buffered saline
PDCoV	Porcine deltacoronavirus
PEDV	Porcine epidemic diarrhoea virus

PFA	Paraformaldehyde
PHEV	Porcine Hemagglutinating Encephalomyelitis Virus
PK15 cell	Porcine kidney-15 cell
PPF	Porcine primary fibroblast
PRCV	Porcine respiratory coronavirus
PRR	Pattern recognition receptors
PRRSV	Porcine reproductive and respiratory syndrome virus
qPCR	Quantitative polymerase chain reaction
RBD	Receptor-Binding Domain
RdRp	RNA-dependent RNA polymerase
RFP	Red Fluorescent Protein
RIG-I	Retinoic acid-inducible gene-I
RLR	RIG-I-like receptor
RNA	Ribonucleic Acid
RNAi	RNA-interference
RNA-seq	RNA-sequencing
RNF	Ring finger protein
RT	Reverse transcription
RTPase	RNA triphosphate
RTC	Replicase-transcriptase complex
S protein	Spike protein
SADS-CoV	Swine acute diarrhoea syndrome coronavirus
SARS-CoV	Severe acute respiratory syndrome coronavirus
sgRNA	Sub-genomic RNA
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
SIV	Simian immunodeficiency virus
SNP	Single nucleotide polymorphism
SOCS	Suppressor of cytokine signalling
ST cell	Swine testis cell
STAT	Signal transducer and activator of transcription

TBK1	TANK-binding kinase 1
TCID50	Tissue culture infectious dose 50%
TGE	Transmissible gastroenteritis
TGEV	Transmissible gastroenteritis virus
TLR	Toll-like receptor
TMPRSS2	Transmembrane protease serine 2
Tom70	Translocases of outer membrane 70
TRAF3	Tumour necrosis factor receptor-associated factor 3
TRS-B	Body transcription regulatory sequences
TRS-L	Leader transcription regulatory sequences
Tyk2	Tyrosine kinase 2
URT	Upper respiratory tract
UTR	Untranslated Region
VSV-G	Vesicular stomatitis virus G glycoprotein
WGA	Wheat germ agglutinin
ZAP	Zinc finger antiviral protein

List of figures and tables

CHAPTER 1

General Introduction

1.1 Coronavirus

1.1.1 History and Classification

Viruses are the smallest of all self-replicating organisms, characterized by their ability to pass through filters that retain the smallest bacteria. The simplest viruses consist of little more than a small segment of nucleic acid protected by a protein shell. Viruses are obligate intracellular parasites that have no metabolism of their own and rely on hijacking the host cellular machinery and subverting it for their own purposes. For that reason, whether viruses are living organisms remains debatable.

In 1971, David Baltimore published a short paper organising viruses into six classes based on their genetic system [1]. Shortly after, an extra class was introduced following the genome characterization of hepatitis B viruses (HBV) as they have a very unique replication system [2]. The Baltimore Classification system, named after its creator, classifies viruses based on their genetic material (deoxyribonucleic acid (DNA) or ribonucleic acid (RNA)), polarity (positive or negative), and stranding (double or single) in one of the seven following groups:

Class I: Double-stranded DNA viruses

Class II: Single-stranded DNA viruses

Class III: Double-stranded RNA viruses

Class IV: Positive-sense single-stranded RNA viruses

Class V: Negative-sense single-stranded RNA viruses

Class VI: Single-stranded RNA viruses with a DNA intermediate

Class VII: Double-stranded DNA viruses with an RNA intermediate

In 2016, The International Committee on Taxonomy of Viruses (ICTV) partially integrated the Baltimore system into the new Linnaean-like ranking hierarchy of virus taxonomy based on evidence that certain groups were descended from common ancestors [3]. For example, the newly proposed realm *Riboviria* includes all RNA viruses (and viruses with an RNA intermediate, e.g. HBV) encoding an RNA-dependent RNA polymerase (RdRp) or RNA-dependent DNA polymerase (also known as reverse transcriptase), corresponding to all viruses in Baltimore classes III, IV, V, VI and VII.

The *Nidovirales* is one of the thirty orders that falls under *Riboviria*. Members of the order are enveloped positive-sense single-stranded RNA (ssRNA) viruses with the largest known RNA genome. They are separated from other RNA viruses by a distinctive genomic organization as well as a unique transcription strategy. The family *Coronaviridae* lies under this order and has three subfamilies including *Orthocoronavirinae*, *Letovirinae* and *Pitovirinae*, which are known to infect mammals and birds, frogs [4] and fish [5] respectively. Replication has been studied in detail only for the *Orthocoronavirinae* as they host some of the most significant viruses to humans [6]. This subfamily consists of four genera: alphacoronavirus, betacoronavirus, gammacoronavirus and deltacoronavirus. Bats are considered to be the original hosts of alpha- and betacoronaviruses and they almost exclusively infect mammals. Birds are proposed to be the original hosts for Gamma- and Deltacoronavirus [7].

The first reported case of coronavirus infection was in 1931 when Schalk and Hawn [8] described cases of avian infectious bronchitis in newborn chicks caused by a filterable

virus (later known as infectious bronchitis virus or IBV). In 1946, Doyle and Hutchings [9] identified a virus as the etiological agent of transmissible gastroenteritis in pigs (later known as transmissible gastroenteritis virus or TGEV) and in 1951, Gledhill and Andrewes [10] isolated a virus from mice causing severe hepatitis and neurological disease (later known as mouse hepatitis virus or MHV). However, it was not until the late 1960s that these viruses, as well as some uncharacterized human respiratory viruses (later assigned as 229E), were revealed to share features of club-shaped spikes that projected from the virion surface under electron microscopy [11]. These spikes give the viral particle a unique appearance resembling the solar corona, prompting the name 'coronavirus' to be officially used for the first time in 1968 [12].

1.1.2 Molecular biology

Genome structure and organization

The coronaviruses have one of the largest genomes of all RNA viruses, ranging from 26 to 32 kb. Unlike any other viruses or organism, the genomes of all positive-sense RNA viruses have to serve as mRNA for the production of protein. The calicivirus utilizes a small terminal protein at the 5' end to recruit ribosomes while the picornavirus and potyvirus contain an internal ribosome entry site (IRES) at the 5' untranslated region (UTR) for cap-independent translation [13]. In contrast, coronavirus genomes have standard eukaryotic 5'-terminal Type I cap structures and 3' polyadenylated tails, which can be directly translated by the host machinery to produce viral proteins [14].

All coronaviruses have a basic set of genes including Replicase, Spike (S), Envelope

(E), Membrane (M) and Nucleocapsid (N) and these genes are in an invariant order of 5'-replicase-S-E-M-N-3'. The replicase is the largest gene, occupying two-thirds of the available coding capacity, and is composed of two Open Reading Frames (ORF) namely ORF1a and ORF1ab. These are the only genes that can be directly translated from the viral genome, encoding non-structural proteins (nsps) with enzymatic functions which are responsible for genome replication and transcription. The other four basic genes encode for structural proteins, where S, E and M are responsible for the building of the exterior structure of the virion, while N binds along the RNA genome in a beads-on-a-string configuration and resides inside the virion [14]. The 5'-most position of the replicase gene is believed to be dictated by the fact that expression of this gene is essential before coronaviruses can progress any further into their infection cycle. The organization of subsequent structural genes appear to be less important, as engineered rearrangement of these genes does not affect virus viability [15].

Apart from the basic genes, accessory genes can be found dispersed in the intergenic intervals of structural genes downstream of the replicase gene, although never between the E and M genes [16]. The protein products of these accessory genes can be either structural or non-structural and can be embedded, partially or entirely, as an alternate reading frame within another accessory or structural gene. For example, the ORF3b gene of severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1) lies partially within the ORF3a gene, while the internal gene of MHV lies entirely within the N gene [17]. Worth noting, identically named accessory genes in coronaviruses from different genera, such as ORF3a in TGEV (alphacoronavirus), SARS-CoV-1 (betacornavirus) and IBV (gammacoronavirus) are generally not related as the naming are simply based on the

transcript number in which they are expressed. It has been proposed that these genes are acquired through ancestral recombination with cellular or viral RNA. One such example is the haemagglutinin esterase (HE) gene exclusive to the betacoronavirus lineage A. It is hypothesised that the HE gene was acquired via horizontal gene transfer from influenza C virus after diverging from the ancestors of other betacoronavirus lineages [18, 19]. While most coronavirus accessory proteins are considered dispensable for replication *in vitro*, many have been linked with virulence and pathogenesis *in vivo*. For example, deletion of TGEV ORF3 and MHV ORF2a, HE, ORF4 and ORF5a result in significant attenuation in their natural hosts respectively [20, 21].

ORF1a/b nsps form the Replicase-Transcriptase Complex (RTC), which is responsible for the replication of genomic RNA and transcription of multiple sub-genomic mRNAs (sgRNA or sg mRNA). The expression of the structural and accessory genes downstream of the replicase occurs via the translation of these sgRNAs. All sgRNAs consist of an identical 5' leader sequence of 70 to 100 nucleotides which is derived from the 5' end of genomic RNA. This is followed by a body sequence representing a portion of the 3' end of the genomic RNA. For example, in addition to the full genome, TGEV produces six additional sgRNA as shown in **Fig 1.1**.

The sgRNAs are structurally similar to the genomic RNA, possessing a 5' cap and a 3' polyadenylated tail which allow recognition by the host translational machinery. Together they form a 3'-coterminal, nested set of mRNAs that is the most distinctive characteristic of the order *Nidovirales*. It is this feature that gives the order its name – *nido* is the Latin word for "nest" [22]. It should be noted that although the longer coronavirus sgRNAs

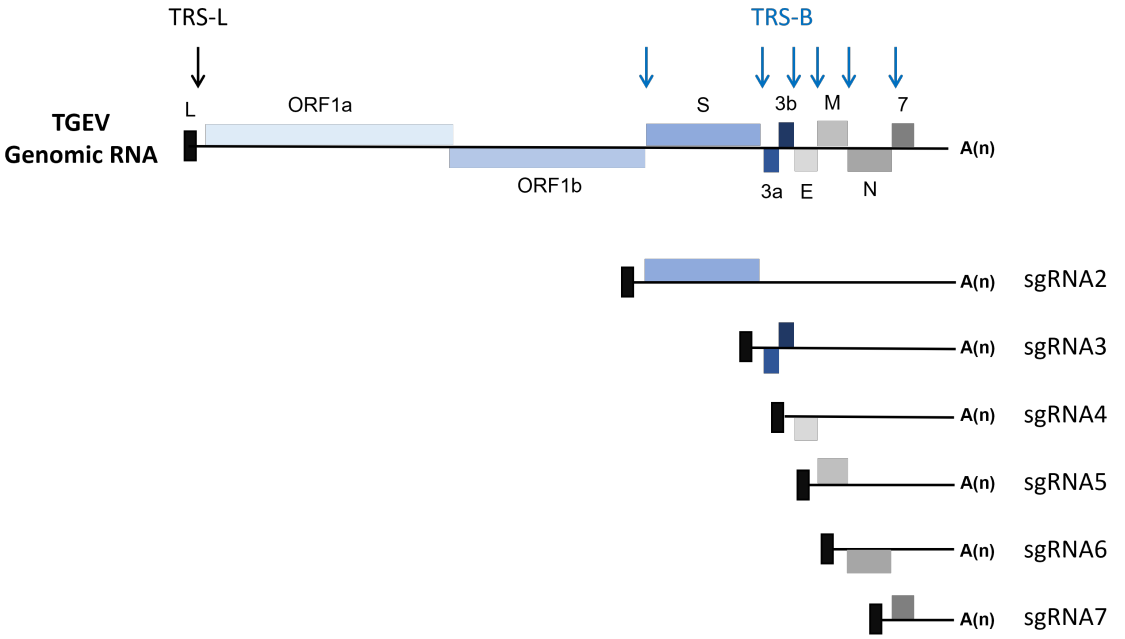


Figure 1.1: **Genome Organization of TGEV**. The single positive-strand genome of transmissible gastroenteritis virus (TGEV) encodes nine open reading frames (ORFs; represented by boxes in different colours and sizes) expressed by the full-length genome and six subgenomic RNA (sgRNAs), which are all 5' capped and 3' polyadenylated. The leader transcription regulatory sequences (TRS-L; black arrow) is located at the 5'end adjacent to the leader sequence while the body TRS (TRS-B; blue arrows) are interspersed at the 5'end of indicated ORFs.

are structurally polycistronic and contain coding sequences for multiple proteins, they are functionally monocistronic. Translation of genes beyond the first ORF at the 5' end does not usually occur, with the exception of some minor accessory proteins such as the ORF3b of TGEV, which is expressed in the same sgRNA as ORF3a [23, 24].

Replication cycle

Coronavirus infection is initiated by the specific binding of S protein to cellular receptors. Different entry receptors have been identified for different coronaviruses. Examples include aminopeptidase N (APN, also known as CD13) for TGEV and angiotensin-converting enzyme 2 (ACE2) for SARS-CoV-1. The differential expression and distribution of entry receptors in different tissues are determinants for viral tropism and

pathogenicity [23]. The coronavirus S proteins are homotrimeric and contain two functionally distinct subunits, S1 and S2. The protruding S1 contains the receptor-binding domain (RBD) while the transmembrane S2 contains the fusion peptide, which is responsible for membrane fusion [25]. The binding of S1 leads to conformational rearrangements, activating S2 which directs the fusion of the viral membrane with either the cell surface or endosome membrane of the target cell.

This fusion leads to the release of the viral nucleocapsid and genome into the host cell cytoplasm and initiation of viral gene expression. The first gene to be translated is the replicase, which results in the expression of two polyproteins, pp1a and pp1ab (**Fig 1.2**). During translation of pp1a, some ribosomes will terminate once they reach the stop codon of ORF1a, while 45 to 70% of them will skip the termination signal and continue their translation through ORF1b, resulting in pp1ab [26, 27]. This mechanism is termed 'ribosome frameshifting', where a translating ribosome shifts one nucleotide backwards from the reading frame of ORF1a into ORF1b [14]. This is controlled by two RNA elements near the overlapping region, a 'slippery' sequence 5'-UUUAAAC-3' which is conserved in all known coronaviruses, along with an RNA pseudoknot structure located downstream of the slippery sequence [28]. This frameshifting determines the stoichiometry between pp1a and pp1ab, producing approximately 1.4 - 2.2 times more of pp1a than pp1ab. These polyproteins then undergo autoproteolytic cleavage into sixteen nsps, except for gammacoronaviruses, which have only fifteen. Autoproteolytic cleavage is driven by the papain-like proteases (PL^{pro}) within nsp3 and the main protease (M^{pro}) nsp5, generating nsp1-11 from pp1a and nsp1-10 and nsp12-16 from pp1ab.

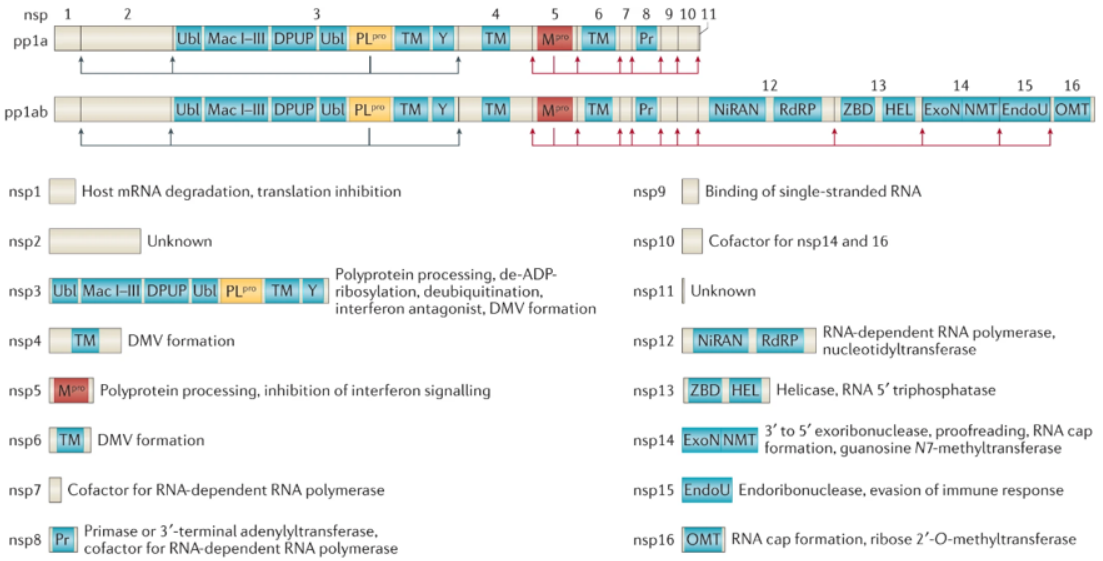


Figure 1.2: **Function of non-structural proteins of SARS-CoV-2**. The ORF1ab of SARS-CoV-2 and other coronaviruses will be translated into polyprotein pp1a and pp1ab. These polyproteins will be cleaved by viral proteases to produce fifteen to sixteen non-structural proteins (nsps) which have diverse functions and are responsible for host immune evasion and virus replication. Adapted from V'kovski et al [23]. Reproduced with permission from Springer Nature, License Number: 5880330574475.

Once the polyproteins have been processed, the coronavirus RTC is assembled from the nsps. Nsp1-11 are involved in producing a suitable environment for viral RNA synthesis, such as evasion of host immune response and rearrangement of intracellular membranes, while nsp12-16 contain the actual enzymes such as RdRp and helicase for RNA replication and transcription [29, 30]. Coronavirus infection induces extensive remodelling of membranes that leads to the formation of replication organelles such as double-membrane vesicles (DMVs) and convoluted membranes (CV), where viral RNA synthesis occurs [31].

Negative-sense RNA has to be synthesized from the positive-sense genome to act as template for the production of genomic RNA. This will then be used for virion packaging, template for more negative-sense RNA synthesis or translation to generate viral proteins. What differentiates coronaviruses and members of the *Nidovirales* from other Group IV

viruses is the production of sgRNA by discontinuous transcription [32]. During synthesis of negative-sense RNA, some RTC will pause when it encounters the body transcription regulatory sequences (TRS-B) located just upstream of the majority of ORFs (**Fig 1.1**). The nascent negative strand (containing the negative-sense TRS-B) will then dissociate and switch templates, binding complementarily to the TRS located adjacent to the leader sequence (TRS-L) on the positive-sense template and continue to transcribe, resulting in a negative-sense sgRNA containing an antileader. Occasionally, the RTC will bypass the TRS-B and continue transcription, until it reaches another TRS-B or transcribes the full-length genome. This discontinuous transcription results in the production of a set of negative-sense sgRNAs that are used as templates for the synthesis of positive-sense sgRNAs. These are translated into structural and accessory proteins [32, 33].

Translated structural proteins, including the membrane-bound S, M and E are initially translocated and inserted into endoplasmic reticulum (ER) membranes. From there, they transit through the ER-to-Golgi intermediate compartment (ERGIC) where virion assembly takes place. The newly produced genomic RNA encapsidated by N protein will bud into the lumen of ERGIC, thereby acquiring a lipid bilayer containing the structural proteins. The newly enveloped virion then exits the cell through the exocytic pathway [14, 23, 34, 35].

1.1.3 Infection in humans and pigs

Coronavirus can be found in a broad range of species, from domestic animals like dogs and cats to marine mammals such as seals and beluga whales. In humans, the majority of coronavirus infections cause respiratory disease, with damage of varying

severity depending on the virus. To date, seven human coronaviruses (HCoVs) have been identified, including HCoV-229E, HCoV-NL63, HCoV-OC43, HCoV-HKU1, Middle East respiratory syndrome coronavirus (MERS-CoV), SARS-CoV-1, and the recently emerged SARS-CoV-2. The first two belong to the genus alphacoronavirus with the rest belonging to the betacoronavirus genus. The first four HCoVs generally cause mild illnesses in the upper respiratory tract and contribute to 15 to 30% of common cold cases in adults. Whereas the other human coronaviruses cause more severe respiratory disease with mortality rates ranging from 2 to 35% [36–38].

Before the outbreak of SARS-CoV-1 in 2002–2003 in Guangdong, China, HCoV-229E and HCoV-OC43 were the only coronaviruses that were known to be circulating in the human population, causing the common cold. They were considered relatively harmless due to the mild illness they cause in immunocompetent individuals. SARS-CoV-1 in contrast, as its name suggests, causes severe acute respiratory infection with morbidity and mortality rates as high as 11% [39]. The first SARS epidemic highlighted the importance and potential threat of coronaviruses, leading to increased research interest and the discovery of many novel coronaviruses in humans and other animals. In 2003 HCoV-NL63 was isolated from a child suffering from bronchiolitis and conjunctivitis, and in 2004 HCoV-HKU1 was isolated from an elderly patient with pneumonia [40, 41]. However, these two coronaviruses did not cause severe disease or large-scale outbreaks. Ten years after the discovery of SARS-CoV-1, another highly pathogenic coronavirus was discovered in Saudi Arabia in 2012. MERS-CoV causes acute pneumonia and lower respiratory tract infection with a high case fatality rate of 35% [42]. Unlike SARS-CoV-1 which disappeared a year after its emergence, MERS-CoV continued to cause

outbreaks in the Middle East and South Korea between 2014 and 2016, probably because it has a more established reseviour host (i.e. camels) before the initial outbreak [43]. In December 2019, another SARS-like coronavirus was discovered and isolated from a patient suffering from pneumonia in Wuhan, China, which soon spread across the world, causing the coronavirus-induced disease 2019 (COVID-19) pandemic [44].

Pigs (*Sus scrofa*) are an important host for coronaviruses. To date, six different coronaviruses have been identified in pigs, with four alphacoronaviruses, including TGEV, porcine respiratory coronavirus (PRCV), porcine epidemic diarrhoea virus (PEDV), swine acute diarrhoea syndrome coronavirus (SADS-CoV); one betacoronavirus - porcine hemagglutinating encephalomyelitis virus (PHEV) and the porcine deltacoronavirus (PDCoV). Among them, TGEV, PRCV, PEDV and PHEV have been circulating in pigs for decades, whereas SADS-CoV and PDCoV are considered newly emerged and had only been identified recently [45, 46]. Unlike coronavirus infections in humans, the majority of porcine coronaviruses are enteric in nature, except PRCV which primarily causes respiratory diseases and PHEV which is neurotropic. The mortality rate of coronavirus diseases in newborn piglets is exceptionally high, ranging from 40 to 100%, with the exception of PRCV which has a negligible mortality [47].

PHEV is the cause of a vomiting and wasting disease and/or encephalomyelitis in pigs. The first clinical outbreak was reported in 1957 in Canada, but the virus was not isolated until five years later [48, 49]. It replicates primarily in the respiratory tract but will spread to the central nervous system through the peripheral nervous system. Usually, PHEV causes subclinical infections in older pigs, while mortality in piglets younger than 4

weeks can be as high as 100% [47]. PHEV is now considered endemic globally, although sporadic outbreaks have been reported [50].

TGEV is the etiological agent of acute swine enteric disease transmissible gastroenteritis, which causes symptoms such as vomiting, diarrhea and severe dehydration in infected animals. The mortality rate of TGEV-infected piglets can be as high as 100% and surviving pigs usually show growth retardation. This caused huge global economic losses in the pork industry until the late 90s [51]. The decline of TGEV is believed to be attributed to the emergence of PRCV, a naturally occurring mutant of TGEV first discovered in 1984 in Belgium [47, 52]. PRCV has a major deletion (621-681 nucleotides depending on strain) in the S gene near the N terminus and variable deletion in the ORF3 gene depending on the strain. Despite the deletion in the S gene, the receptor-binding domain of PRCV does not seem to be affected and it binds like TGEV to the same host receptor, APN [53]. However, the deletion does render PRCV unable to bind sialic acids, which is thought to play an important role in TGEV enteric tropism and thus partially explains the shifted respiratory tropism of PRCV [54]. The S protein and ORF3 of TGEV have also been shown to play a critical role in virus virulence, as some TGEV isolates with minor deletions (6 nucleotides in the case of Purdue P115) have an attenuated phenotype. This suggests the milder disease caused by PRCV may also be attributed to deletions in the S and ORF3 genes [20, 51, 55]. Antibodies produced as a result of PRCV infection cross-react with TGEV and are therefore protective. PRCV spread quickly in Europe and naturally immunized the pig population without causing significant health problems, leading to the elimination of transmissible gastroenteritis, with TGEV no longer a significant problem for pig farmers in Europe [47]. However,

sporadic outbreaks have still been reported in China [56].

Other emerging and re-emerging porcine enteric coronaviruses have replaced TGEV as the major threat to the global pig farming industry. PEDV was identified in the 70s and spread across Europe in the following 20 years. Although significant outbreaks have not been reported recently in Europe, outbreaks are observed in Asia and the diseases are more acute and severe. PEDV remains the second most important porcine pathogen in China. In fact, emerging cases caused by a highly virulent strain have been found across the world since 2013, raising concerns of re-emergence [57]. PDCoV was identified in 2009 in Hong Kong but its role in porcine enteric disease was not known until it caused diarrhea in pigs in the US in 2014 [7, 45]. Since then, PDCoV has become one of the major causes of porcine enteric disease in China and the US. Co-infection of PDCoV and PEDV was found to be common in China, underlying the risk of virus recombination [58]. SADS-CoV emerged in China in 2016 and the outbreak in 2017 and 2019 in southern China has caused the death of tens of thousands of piglets and huge economic loss [46, 59]. As yet, SADS-CoV has not been reported outside of China.

1.2 Interplay between viruses and innate immunity

The innate immune system is essential for the host's protection from infectious diseases and is the first line of defence against viral infections, limiting different steps in the virus life cycle including entry, translation, replication and assembly. The presence of pathogens can be directly detected by pattern recognition receptors (PRRs) which recognize unique structural features of microorganisms called pathogen-associated molecular patterns (PAMPs), or indirectly via the alteration of normal cellular processes by viruses such as inhibition of translation and ER stress. The major targets for the recognition of viruses are their nucleic acids, as the structural features of viral RNA or DNA are distinct from the host's. These include double-stranded and 5'-triphosphate viral RNA, as well as viral DNA with unmethylated CpG motifs. PRRs are usually located in endolysosomes and the cytosol, as viral but not host nucleic acids are often found in these locations, aiding the differentiation of foreign and self molecules. Activation of these sensors will lead to the production of interferons (IFN) and other pro-inflammatory cytokines (discussed in detail later using SARS-CoV-2 as an example), as well as the development of adaptive immunity [14, 60].

1.2.1 Host Interferon response

Multiple PRRs exist, each playing a specific role in pathogen detection [61]. Cytoplasmic DNA (both viral and self-derived) is detected by sensors such as AIM2 and cyclic GMP-AMP synthetase (cGAS), while cytoplasmic RNA (single and double-stranded)

is detected by retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs) such as RIG-I, melanoma differentiation-associated gene 5 (MDA5) and laboratory of genetics and physiology 2 (LGP2) [62]. RLRs recognise RNA through their helicase domains, catalyzing a conformational change and translocation, resulting in interaction with a common downstream adaptor, the mitochondria antiviral signalling protein (MAVS) [63], which localizes in mitochondrial membranes and peroxisomes. The recruitment of tumour necrosis factor receptor-associated factor 3 (TRAF3), TANK-binding kinase 1 (TBK1) and IκB kinase epsilon (IKKε), followed by MAVS activation, leads to the phosphorylation and dimerisation of the constitutively expressed major IFN gene regulator IFN regulatory factor 3 (IRF3) in the cytoplasm [64]. Another major IFN gene regulator is IRF7, which is expressed following initial IFN expression (i.e. IFNβ and IFNα4), can form heterodimers with IRF3 and is critically involved in later response [65]. These homo- and heterodimers will then translocate into the nucleus, triggering the expression of IFNs (**Fig1.3**). The activation of PRRs will also trigger the translocation of other transcription factors such as nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), which is responsible for producing pro-inflammatory cytokines.

To date, IFNs can be classified into type I (IFNα, IFNβ, IFNε, IFNκ, and IFNω), II (IFNγ), and III (IFNλs) in humans. Type I and III IFNs bind to IFN-alpha receptor 1 and 2 (IFNAR1/2) heterodimers and IFN-lambda receptor 1 (IFNLR1) and interleukin-10 receptor 2 (IL10R2) heterodimers respectively on the cell surface. By autocrine and paracrine effects, they activate Janus kinase 1 (JAK1) and tyrosine kinase 2 (Tyk2), which in turn lead to the phosphorylation of the signal transducer and activator of transcription (STAT) 1 and 2 [66]. Heterodimers formed by phosphorylated STAT1/2

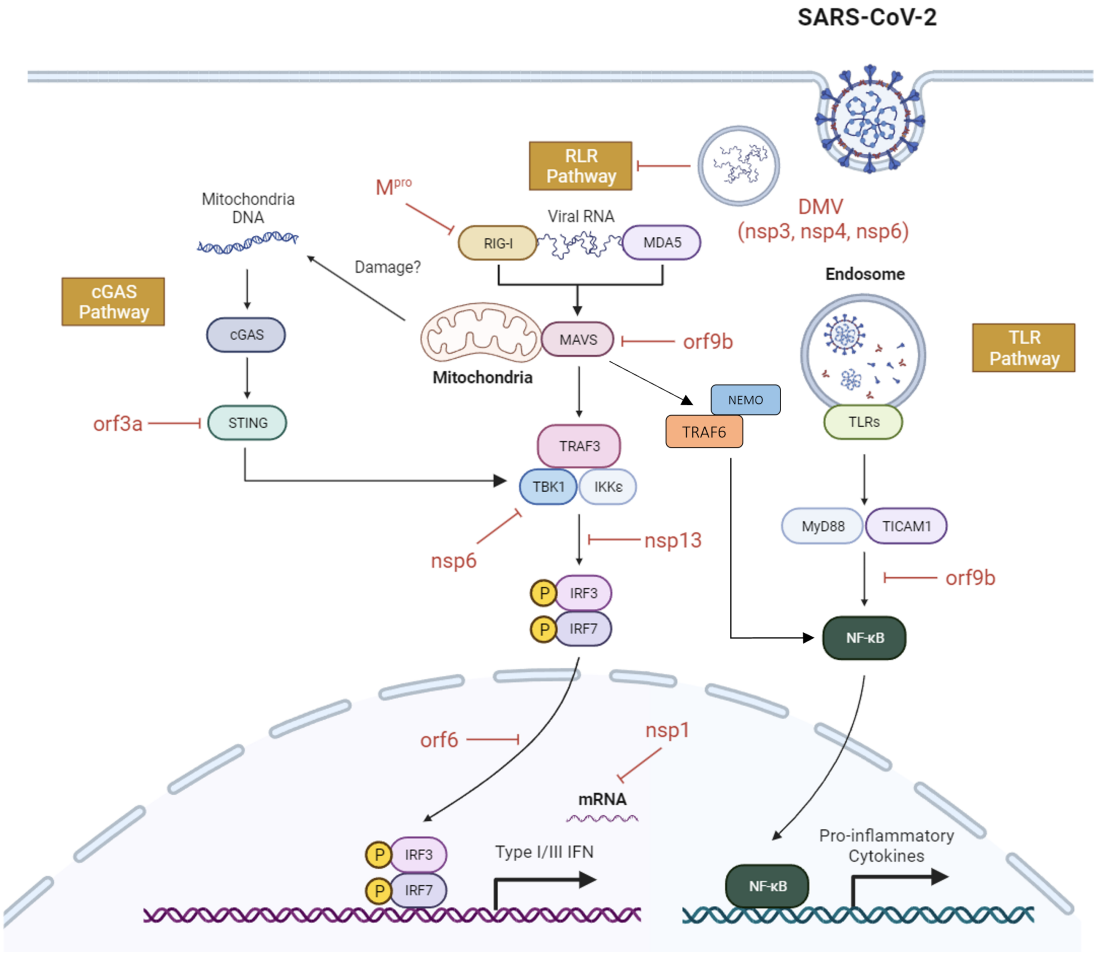


Figure 1.3: **Innate immune sensing and evasion of SARS-CoV-2.** SARS-CoV-2 infection can trigger multiple pathways in the innate immune system to raise an antiviral response. The single-strand viral RNA is primarily detected by the cytoplasmic RNA sensors retinoic acid-inducible gene I (RIG-I) and melanoma differentiation-associated gene 5 (MDA5) to induce interferon (IFN) production and other pro-inflammatory cytokines via nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB). Viral infection may cause mitochondrial damage, leading to mitochondrial DNA release and activation of cGAS. Other pathogen-associated molecular patterns (PAMPs) such as the spike and envelope protein of SARS-CoV-2 may be sensed by toll-like receptors (TLRs), activating the NF-κB pathway and production of pro-inflammatory cytokines. However, viral proteins encoded by SARS-CoV-2 (in red) can interfere with these pathways and evade the host innate immune response. (Created with BioRender.com).

recruit IRF9, forming a complex known as IFN-stimulated growth factor 3 (ISGF3). Type II IFN binds to IFNGR1 and IFNGR2 heterodimers, leading to the phosphorylation of preassociated JAK1 and JAK2 tyrosine kinases, resulting in the recruitment and phosphorylation of STAT1. Phosphorylated STAT1 homodimers then form the IFNγ activation factor (GAF). Both ISGF3 and GAF translocate into the nucleus and bind to interferon-stimulated response elements (ISRE) and gamma-activated sequence (GAS) promoter elements respectively, leading to the expression of interferon-stimulated genes (ISGs) (**Fig 1.4**). Type I IFNs are the most extensively studied among the three, and are thought to be responsible for the majority of antiviral activities. Less is known about type III IFNs, but their production following PRR activation and association with the JAK-STAT signalling pathways similar to type I IFNs suggest they are also involved in antiviral activities. In contrast, type II IFNs are involved in inflammation, as well as priming the activation of adaptive immune response [67–69].

Interferon-stimulated genes

ISGs are generally defined as genes that are upregulated following IFN stimulation and their counterparts are interferon-repressed genes (IRGs), which limit the IFN response avoiding tissue damage and pathology associated with an overactive response. Triggering the IFN response ultimately results in the activation of hundreds of ISGs, creating a pleiotropic and redundant antipathogenic state within the stimulated cell. ISGs that can interfere with the entry, replication, egress, or other stages of the viral life cycle are referred to as antiviral factors or restriction factors and are the actual effectors contributing to innate antiviral immunity [68]. Understanding the mechanisms of these direct-acting antiviral factors has been an active field of research as this deepens our

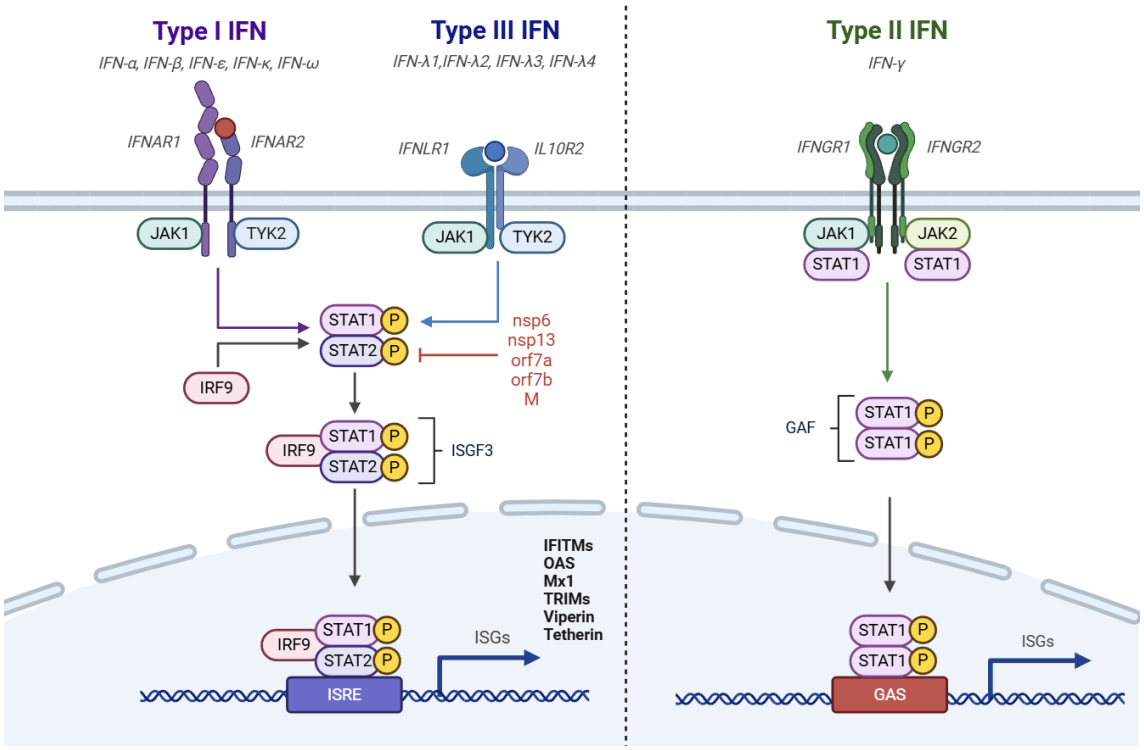


Figure 1.4: **Signalling pathways of IFN and SARS-CoV-2 inhibition** The three types of interferons signal through specific cell surface receptors (IFNAR1/INFAR2 for type I, IFNGR1/IFNGR2 for type II and IFNLR1/IL10R2 for type III) to induce transcription of interferon-stimulated genes (ISGs). Receptor engagement leads to activation of intracellular kinases (JAK1, JAK2, TYK2) that phosphorylate STAT1 and STAT2. The activated transcription factors (ISGF3 for type I/III or GAF for type II) translocate to the nucleus to induce transcription by binding ISRE or GAS promoter elements. SARS-CoV-2 encoded viral proteins (in red) can prevent ISG transcription by blocking phosphorylation and nuclear translocation of said factors. STAT: signal transducer and activator of transcription; ISGF3: interferon-stimulated growth factor 3; GAF: interferon-gamma-activated factor; ISRE: interferon-stimulated responsive element; GAS: gamma-activated sequence (Created with BioRender.com).

understanding of virus-host interactions.

One could imagine that the most promising antiviral strategy for the host is to prevent the virus from infecting the cell in the first place. Indeed, there are multiple examples of ISGs targeting viral entry. The interferon-induced transmembrane (IFITM) proteins may be the best-studied of these. The three best-characterized members are IFITM1, 2 and 3, with IFITM1 localized at the plasma membrane and IFITM2 and 3 on endo-lysosomal membranes [70]. These proteins were first shown to be potent inhibitors of influenza A virus (IAV) and have since been shown to inhibit a diverse range of viruses through their palmitoylated amphipathic alpha helix, which blocks fusion between the virus envelope and cellular membrane [71–73]. Another example is the Cholesterol-25-hydroxylase (CH25H), an IFN inducible enzyme that converts cholesterol into 25-hydroxycholesterol (25HC). This oxysteroid is likely responsible for the antiviral effect as treating cells with 25HC or transferring supernatants from CH25H-expressing cells convert protection against enveloped viruses [74]. It has been proposed that 25HC alters membrane properties, probably by controlling cholesterol trafficking, to suppress viral entry [68]. The human nuclear receptor coactivator 7 (NCOA7) is another more recently identified and characterized entry inhibitor [75, 76]. It appears that NCOA7 specifically restricts the entry of viruses that use the endocytic route, binding to the H+-ATPase in vacuoles and promoting vesicle acidification and degradation of endocytosed cargo [77].

Other ISGs such as the tripartite motif-containing protein 5, alpha isoform (TRIM5α), and myxovirus resistance protein 1 and 2 (Mx1 and Mx2) target the post-entry step to

restrict viruses from reaching their final destination, such as the human immunodeficiency virus (HIV) 1 and IAV which need to enter the cell nucleus for the initiation of their replication cycle. For example, TRIM5α can form restrictive hexameric cages around the HIV capsids, which inhibits viral uncoating and nuclear entry [78–80]. The Mx proteins work in a similar principle. Mx1 is believed to surround released viral nucleocapsids of IAV and direct them to degradation, whereas Mx2 acts by blocking the HIV-1 viral genome from entering the nucleus [81].

After entering the cell, all viruses regardless of their genetic system, must produce mRNA and protein to continue their life cycle. The IFN system has equipped cells with ISGs that target viral gene transcription, genome replication, and protein translation. Unlike RNA viruses which code for their own RdRp, DNA viruses and retroviruses hijack the host DNA-dependent RNA polymerase for gene expression. Inhibition of transcription of viral DNA is deleterious as this will almost inevitably affect the host's own gene expression system. Nonetheless, the dsDNA sensor IFN gamma inducible protein 16 (IFI16) was found to impair the transcription of two genes in human cytomegalovirus (HCMV) responsible for viral DNA synthesis [82]. In contrast, transcription from RNA (by RdRp or reverse transcriptase) is a more desirable target as this is a unique feature of RNA viruses. RBBP6 mimics the viral nucleoprotein and blocks Ebola virus transcription and replication; SAMHD1 decreases intracellular deoxynucleotide triphosphate (dNTP) level to prevent reverse transcription and APOBEC3G deaminates cytosines to hypermutate reverse-transcribed viral DNA of retroviruses and HBV; Viperin (or RSAD2) can restrict viruses through multiple mechanisms, such as catalysing the formation of a nucleotide analogue that acts as a chain terminator for viral RNA extension by the RdRp of

flaviviruses [83–86].

Eukaryotic and viral mRNA differ in many aspects such as chemical modification, nucleotide bias, and conformation. The host immune response can discriminate self from non-self mRNA based on these differences and target viral mRNA for degradation or prevent them from being translated [87]. For example, host mRNA will undergo post-transcriptional modification to generate a type 1 cap. The IFN-induced protein with tetratricopeptide repeats (IFIT) can recognize foreign mRNA without caps (i.e. IFIT1) or with a type 0 cap lacking 2'-O-methyl modification (i.e. IFIT2) and bind to them, preventing them from being translated [88–90]. Another example is the Zinc finger antiviral protein (ZAP). Plant and vertebrate genomes have significantly lower CpG dinucleotide frequencies than expected, meaning the expected number of cytosine-guanine pairs is low relative to the overall GC content in their genome [91, 92]. ZAP can recognise and directly bind to viral transcripts with high CpG levels and recruit the RNA-processing exosome to degrade them [93, 94]. Some RNA secondary structures are unique to pathogens and can be recognized as non-self RNA by the host. One such example is long dsRNA in the cytosol as this is a hallmark of virus replication and is not usually found in healthy uninfected cells [87]. Apart from the PRRs such as RIG-I and MDA5 which induce IFN production, some sensors can, upon recognition, directly inhibit viral replication. The 2'-5'-oligoadenylate synthetase 1 (OAS1), when bound to dsRNA, will activate ribonuclease L (RNase L), which degrades the targeted RNA [95, 96]. The adenosine deaminase acting on RNA 1 (ADAR1) catalyzes the hydrolytic deamination of adenosine to inosine in dsRNA, causing a mutation from adenosine to guanosine (inosine is read as guanosine by ribosomes), which may lead to non-synonymous mutation and

potentially the translation of non-functional viral proteins [97].

The last stage of the virus replication cycle is assembly and egress. Interestingly, at least to date, very few ISGs are known to target this phase. Among them, the most studied ISG is tetherin (encoded by the gene *BST2*), a transmembrane protein first described to trap HIV-1 virions on the plasma membrane [98]. It uses two membrane anchors to trap budding virions on the cell surface and prevent them from being released. Since its discovery, this antiviral mechanism has been found to extend to several other enveloped viruses [68, 76].

Additional interferon-stimulated response

Apart from the classical definition of IFN response, where IFN triggers the transcription of ISGs, changing the cellular proteome to establish an antiviral state within the cell, recent studies utilizing advanced sequencing and proteomic technologies have given us new insights into the additional effect of IFN stimulation [68]. Some ISGs can yield multiple isoforms by alternative splicing and give rise to proteins with different cellular localization and activity. For example, ADAR1 is constitutively expressed in both the nucleus and cytoplasm, and upon IFN stimulation, an alternative transcript is produced that gives rise to an extended protein which predominately localized to the cytoplasm [99]. ZAP has multiple isoforms and while the long isoform is constitutively expressed, the short isoform is only induced by IFN. Despite all isoforms being antiviral, they were shown to have differential potency against different viruses [100]. It is, therefore, possible that IFN can promote splice variants that are otherwise not antiviral. These changes in isoform preference might not be detected by conventional methods which only measure

changes in the general mRNA abundance of a gene, especially those that differ only in their untranslated regions and are underrepresented during IFN response. Similarly, IFN can also promote the production of specific protein isoforms, for example, by controlling post-translation modifications such as phosphorylation and polyubiquitination as seen in PRRs [101].

The need for speed is crucial to a successful antiviral response as pathogens have numerous ways to counteract host defence mechanisms (discussed in the next section). Having selected mRNAs ready for translation prior to ISG transcription is an efficient way to ensure the rapid production of selected proteins which can enhance pathogen sensing or amplify the IFN response [68]. The mRNA of IFN-β has been reported to pre-exist in uninfected cells but with its translational competence and stability impaired [102]. Type I IFN signalling will lead to rapid changes in translation regulation, influencing the mRNA stability, and enhancing the translation and production of IFN-β. Several other examples suggest that some genes might be constitutively expressed at the mRNA level but not the protein level, and instead of triggering transcription, IFN triggers their translation [103]. These genes will not be captured by transcriptomics analysis such as RNA-sequencing (RNA-seq) and are usually not classified as ISGs. This hypothesis is further strengthened by the fact that discrepancies are usually found between transcriptomics and proteomics studies regarding IFN response.

Innate immunity and SARS-CoV-2

Innate immunity has been shown to play a critical role in the control of SARS-CoV-2 infection clinically, as loss-of-function mutation or carrying certain haplotypes of

components in the IFN response have been associated with severe COVID-19 [104–106]. RNA from genomic, subgenomic, or replicative intermediates of SARS-CoV-2 can be detected by cellular RNA sensors, including MDA5 and RIG-I, which triggers the IFN response. Apart from the RLRs, TLRs are also involved in the sensing of SARS-CoV-2. Multiple studies suggested the S and E proteins can act as ligands and bind to these TLRs, triggering the NF-κB and mitogen-activated protein kinase (MAPK) pathway through myeloid differentiation primary response 88 (MyD88) or TIR-domain containing adaptor molecule 1 (TICAM1, also known as TRIF) to induce transcription of pro-inflammatory cytokines such as interleukins and tumour necrosis factors (**Fig1.3**) [60]. These pathways have also been shown to play a role in the IFN response, as X chromosomal genetic anomalies in TLR7 have been linked to severe COVID-19 in young adults [104]. Mutation in other factors involved in IFN signalling such as TBK1, IRF3, IRF7, and IFNAR1/2 have also been linked to severe symptoms [106].

SARS-CoV-2 is sensitive to IFN and the activity of ISGs was found to inhibit their replication [107, 108]. Lymphocyte antigen 6 complex, locus E (LY6E) was found to enhance cellular infection of viruses including IAV and several flaviviruses and alphaviruses [109, 110]. However, this ISG was found to inhibit the entry of a wide range of human coronaviruses including HCoV-229E, HCoV-OC43, MERS-CoV, SARS-CoV-1 and 2, and even the mouse coronavirus MHV by interfering with spike protein-mediated membrane fusion [111]. Strong evidence suggests LY6E protects secretory club and ciliated cells in the respiratory tract from SARS-CoV-2 infection and prevents damage in the lung *in vivo* [112]. As mentioned above, OAS1 pairs with RNase L to recognize and cleave dsRNA. However, a study found that only a certain isoform of OAS1 can effectively targets SARS-

CoV-2. Alleles containing guanine at Rs10774671 can produce a C-terminally prenylated form of OAS1, which is essential for directing them to perinuclear structures rich in viral dsRNA such as the replication organelles or DMV found in SARS-CoV-2 infected cells. Meanwhile, non-prenylated OAS1 is diffusely localized across the cell and ineffective at inhibiting SARS-CoV-2. Analysis of OAS1 transcripts in hospitalized COVID-19 patients shows that this single nucleotide polymorphism (SNP) is associated with protection from severe diseases [105]. A gain of function screen including 399 ISGs has also revealed that a wide range of ISGs can inhibit discrete replication steps of SARS-CoV-2, including previously mentioned IFITMs, IFITs and tetherin [107].

However, IFN (notably IFNγ) can also exacerbate disease by leading to excess production of pro-inflammatory cytokines which causes widespread inflammation, referred to as a 'cytokine storm', especially during later stages of SARS-CoV-2 infection [113]. These cytokines can induce a form of inflammatory cell death termed PANoptosis. Cells that die from PANoptosis release more cytokines, leading to a positive feedback loop that ultimately causes systemic inflammation, multiorgan failure, and lethality [60]. Severe COVID-19 patients were shown to have elevated cytokines with low levels of IFN-I in the early stage and elevated levels of IFN-I during the advanced stage of the disease [114]. This is probably due to a delayed or impaired IFN response in infected cells which leads to increased viral load and cell death. This promotes the accumulation of immune cells including dendritic cells and monocytes at the site of infection, resulting in lung immunopathology and the release of more pro-inflammatory cytokines [115]. The outcome of IFN-I treatment is therefore, perhaps not surprisingly, determined by timing, as it is only protective when administered before or early in the course of infection and

late administration could be detrimental [116].

1.2.2 Innate immune evasion of viruses

The red queen hypothesis proposed in 1973 suggests that species must constantly adapt and evolve to survive in the face of evolving enemies [117]. Viruses and their hosts are also constantly evolving, despite the evolution of a wide variety of 'weapons' in the antiviral arsenal, throughout time, viruses have also developed factors to break through these defences and persist. Similar to how different ISGs target different stages of the viral life cycle, viruses encode virulence factors to counteract different aspects of the innate immune response with multiple strategies, from general sensing to specific antiviral effectors, with the ultimate goal of manipulating the host environment for their own benefit [118].

Evasion from pattern recognition receptors

As described in the previous section, multiple types of PRRs are localized in different compartments within the cell, ready to detect the presence of foreign molecules. RNA viruses that replicate in the cytoplasm, such as flaviviruses and coronaviruses will generate replication intermediates expected to be detected by the cytoplasmic RIG-I and MDA5 to induce IFN. However, infected cells were found to have their cytoplasmic membrane heavily remodelled, creating structures such as DMVs. These replication organelles are the site of RNA replication and are isolated from the cytoplasm (**Fig 1.3**). Transportation of materials is controlled by a pore complex and therefore the aforementioned sensors cannot sense the presence of dsRNA [34, 119]. Hepatitis C virus (HCV), has

been found to hijack the host nuclear pore complex proteins, incorporating them into the membrane of their replication complex to control protein import. Intriguingly, artificially adding a nuclear localization signal to RIG-I overcomes their restricted access to the replication compartments, enabling dsRNA sensing and induction of IFN signalling [120]. For coronavirus, the co-expression of nsp3, 4 and 6 is sufficient to induce the formation of DMVs, with nsp3 being the major component of the pore complex [121–123]. How exactly these pores selectively allow entry of the coronavirus RTC while blocking the entry of PRRs remains unknown.

Viruses also evade host sensors by encoding proteins that protect or mask the viral genome. During HIV replication, the RNA genome is reverse-transcribed into complementary DNA (cDNA) before being transported into the nucleus for integration. The capsid of HIV-1, but not HIV-2 can interact with the host protein cyclophilin A, stabilising their capsid and preventing the viral cDNA from being sensed by the cytoplasmic DNA sensor cGAS [124]. This likely contributes to the different pathogenicity between the two viruses, explaining why HIV-1 is more prominent in the human population.

Some viruses apply a more proactive approach, instead of hiding from the sensors, they encode proteins to target PRRs. Similar to coronavirus, poliovirus is a cytoplasmic RNA virus targeted by RIG-I and MDA5 and encodes proteases to process their polyprotein. In addition to polyprotein cleavage, the two proteases encoded by picornaviruses $2A^{pro}$ and $3C^{pro}$ have been shown to cleave MDA5 and RIG-I respectively, while the M^{pro} of SARS-CoV-2 has been shown to cleave off the N-terminal of RIG-I [125, 126]. Instead of direct cleavage of the sensors, the PL^{pro} of SARS-CoV-2 antagonises the ISGylation of

MDA5, which is required for activation[127]. In addition to targeting PRRs, viral proteins also target downstream signalling molecules in the antiviral innate immune pathways to suppress IFN production.

Inhibition of Type-I interferon response

In addition to cleaving RIG-I, the M^{pro} of SARS-CoV-2 has also been shown to promote ubiquitination of MAVS and target them for proteosome-mediated degradation [126]. Many viral proteins of SARS-CoV-2, both structural and non-structural, were found to play a role in suppressing IFN-I production [128]. An unbiased screen identifying SARS-CoV-2 IFN-I antagonists found that nsp6 can inhibit IRF3 phosphorylation by binding to TBK1, nsp13 can bind and block phosphorylation of TBK1, and ORF6 can bind to an importin to prevent IRF3 nuclear translocation. While nsp1, 6 and 13, ORF3a, 7a and 7b, as well as the M protein, were shown to block STAT1 and 2 phosphorylation and nuclear translocation [129]. Other viruses such as picornavirus and flavivirus also encode for proteins with similar functions that degrade or block these signalling molecules [119]. Blocking the nuclear translocation of transcription factors such as IRF3 and 7, as well as components of the ISGF3 complex is an effective way to inhibit innate immune response as they are responsible for the transcription of IFN-I and ISGs (**Fig 1.3 and 1.4**).

While MAVS is the adaptor for RNA sensors, STING is the adaptor for DNA sensors such as cGAS and IFI16. Perhaps as expected, DNA viruses such as human adenovirus, herpesvirus and papillomavirus all code for proteins that target STING [130, 131]. Interestingly, coronaviruses and flaviviruses also code for proteins to counteract STING. The fact that DNA and RNA viruses both target this adaptor molecule suggests it has

a critical role in IFN induction irrespective of the nature of the viruses [132]. While the mechanism of how STING plays a role in response to RNA viruses remains unclear, it has been suggested the mitochondrial DNA released as a result of damage from infection may prime the innate immune response [133]. One common strategy used by both DNA and RNA viruses to redirect resources toward viral gene expression and dampen the IFN response is to target host mRNA. Host shut-off of cellular gene expression inhibits the translation of IFN and other immune-related genes [35]. In coronaviruses, nsp1 restricts host translation via mRNA cleavage. Although it has no established RNase activity, the expression of nsp1 causes cleavage of mRNA, including IFNβ, at the 5' end, potentially by recruiting a host endonuclease. Selective shut-off is achieved as the 5'UTR of coronavirus viral transcripts is highly structured and is therefore resistant to nsp1-induced cleavage [134]. IAV encodes endoribonuclease PA-X that selectively degrades spliced mRNA transcribed by the host RNA polymerase II, thus sparing viral transcripts made by the viral RdRp [135, 136]. Herpes simplex virus is also notorious for inducing strong host shut-off through tegument proteins. The prepackaged VHS protein can act immediately after virus entry and target host mRNA made in response to early infection for degradation [137]. Once the virus takes control of the cell, the late proteins can neutralize the RNase activities and allow effective translation of viral mRNA [138, 139].

Counteracting antiviral ISGs

For efficient replication, viruses must overcome redundant cellular defence mechanisms. Despite the small genome size, HIV is highly effective at evading ISGs. APOBEC3C can be incorporated into the virion and cause mutations as previously mentioned. The

Vif protein from HIV-1 can inhibit the packaging of APOBEC3G by inducing polyubiquitination and proteasomal degradation [140]. Vpu which induces detachment of the HIV-1 virion from the cell membrane can interact with tetherin, reducing expression at the surface and its ability to inhibit virus budding [141, 142]. Vpx, an accessory protein found in HIV-2 but not HIV-1, can target the dNTP sequestering SAMHD1 for proteasomal degradation [143]. Although not commonly regarded as an ISG, SERINC5 was suggested to enhance IFN-I expression and depends on IFN-I for its antiviral activity [144]. It localizes to the plasma membrane and is incorporated into budding HIV-1 virions, impairing their subsequent entry. HIV-1 Nef protein can redirect SERINC5 into other endosomal compartments, thereby excluding them from being incorporated into virions[145]. The genome of HIV and many vertebrate RNA viruses are suppressed in CpG dinucleotides, a genotype acquired throughout the co-evolution of their host, to evade ZAP which targets RNA with high CpG dinucleotide content [146].

In some cases, viruses can even subvert ISGs to facilitate replication. The aforementioned endo-lysosomal IFITM3 can restrict the entry of multiple classes of enveloped viruses. However, their effect on coronaviruses is less clear-cut, with contradicting reports suggesting it may inhibit or enhance coronavirus replication depending on cellular context [147, 148]. Nonetheless, HCoV-OC43 is an exception, with consistent results suggesting that virus infection is promoted by all three types of IFN, probably by hijacking IFITM3 for efficient entry [149]. The beta-herpesvirus HCMV was also found to exploit two ISGs to facilitate infection. The vMIA protein encoded by HCMV can relocalise viperin from the endoplasmic reticulum to the mitochondria, thereby reducing cellular ATP generation and disrupting the actin cytoskeleton, facilitating viral replication [150]. Meanwhile, the

detachment inhibitor tetherin was found to enhance HCMV replication by enhancing viral entry [151].

As discussed, the capping of eukaryotic mRNA is an essential modification for efficient translation and serves as a marker of 'self' RNA. The complete capping process requires the activities of at least four enzymes, an RNA triphosphate (RTPase) to remove the γ-phosphate from the nascent mRNA; a guanylyltransferase (GTase) to add the guanosine monophosphate to the dephosphorylated mRNA; an N7-methyltransferase (N7-MTase) to methylate the N7 of the added guanine base (type 0 cap); and a 2'-O-methyltransferase (2'-O-MTase) to methylate the ribose 2'-O of the first nucleotide of the capped mRNA (type 1 cap) [90]. The IFIT family will bind to uncapped or partially capped foreign RNA and prevent them from being translated. Many viruses have evolved ways to cap their RNA and mimic host mRNA to evade IFIT recognition. While Lassa fever virus and IAV snatch caps from the host mRNA, coronavirus and flavivirus encode proteins that can perform mRNA capping [119]. For example, the SARS-CoV-2 nsp13 has an RTPase activity, while the nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain of nsp12 was recently suggested to have a GTase activity. N7-MTase and 2'-O-MTase activities were also found in nsp14 and nsp16 respectively [34]. This suggests that, in principle, SARS-CoV-2 has all the required enzymatic activities encoded in their proteins to generate a type 1 cap, and may contribute to evading the sensing of proteins from the IFIT family [152].

In this section, the interplay between viruses and the host innate immunity has been discussed in detail. While the hosts have derived measures (i.e. antiviral ISGs) to

suppress viruses, viruses have also evolved countermeasures to thrive in their hosts. The outcome of their arms races is never one-sided, as they are determined by multiple external factors. For example, some antiviral ISGs described earlier were identified through artificial overexpression experiments. They may not have an effect *in vivo* as the physiological expression level of such ISG may not be sufficient to exert an antiviral effect against a specific virus. Apart from dosage (e.g. antiviral factors overwhelm viruses by quantity or vice versa), other factors including individual (e.g. genetic variation such as SNP within a species), temporal (e.g. virus suppresses the host before host immune system can be activated or vice versa), and spatial (e.g. specific antiviral factor is not expressed in the virus-targeted cell type or tissue) can also play a role in determining the outcome.

1.3 Cross-species transmission of viruses

Nearly two-thirds of human infectious diseases arise from pathogens shared with wild or domestic animals [153]. In rare occasions, viruses may manage to cause an infection when they come in contact with a new host, resulting in cross-species transmission. However, the increase in the global human population, international travel, and ecological changes have enhanced opportunities for cross-species transmission of known and novel viruses. Changes in agricultural practices and interactions between wildlife, domestic species, and humans have also increased the risk of cross-species transmissions [154]. Most of these transmissions are 'dead ends', with infection occurring in limited individuals with no subsequent transmissions. Occasionally, viruses may acquire random mutations, enabling efficient replication and transmission in the new hosts, thereby causing large-scale outbreaks [155]. The terminology of these transmissions differs depending on the direction of flow. Zoonosis refers to the transmission of a pathogen from an animal to a human, while reverse zoonosis refers to transmission from humans to other animals. The term for transmission from one animal species to another is however loosely defined, with some referring to this as epizootic and others as epizoonosis. Confusingly, both these terms have different meanings, with epizootic originally referring to disease events in non-human populations analogous to an epidemic in humans, while epizoonosis refers to skin diseases caused by parasites such as insects. Transmission of diseases among animals, especially wild animals, usually goes undetected and therefore receives much less attention. Not until relatively recently did the scientific community consider these animal-to-animal transmissions (i.e. intermediate hosts) may act as stepping stones for

viruses to acquire mutations essential for animal-to-human transmission. A more defined term for cross-species transmission among animals is therefore needed. In light of this, the term zoonosis will be used to describe both animal-to-human and animal-to-animal transmission hereafter.

1.3.1 Barriers against zoonotic spillover

Multiple viral outbreaks have been recorded in the last couple of decades in the human population, including HIV, IAV, Ebola virus, SARS-CoV, MERS-CoV and most recently SARS-CoV-2. All these viruses have a zoonotic origin, suggesting the risk from animal viruses should not be overlooked. However, compared to the number of animals and viruses present in the wild, these spillover events are relatively rare. This is because zoonotic transmission of viruses requires multiple events to coincidentally happen at a given point in space and time. Factors influencing the probability of transmission include disease dynamics in the reservoir host, pathogen exposure to potential new hosts, and the species-specific factors that affect host susceptibility to infections. Multiple barriers exist throughout the flow of a pathogen from its original host to a recipient host and large-scale zoonotic outbreaks require the pathogen to overcome all of these barriers. Meanwhile, certain human activities and environmental changes (discussed below) alter these naturally occurring barriers, increasing the risk of zoonotic spillovers.

Pathogen Pressure and Exposure

Pathogen pressure can be influenced by many factors, such as the population of the natural hosts, the amount of pathogens they carry, and the nature of the pathogens them-

selves. For example, an increase in the rodent population in 1991-1992 was hypothesized to have caused zoonotic outbreaks of hantavirus in humans between 1993-1994 in the United States [156]. Meanwhile, the viral loads and excretion rates in the salivary glands of rabies-infected animals and gastrointestinal tracts of wild birds carrying influenza virus determine the transmissibility to potential new hosts [157, 158]. Respiratory viruses that have a shorter infectious window may require a closer interaction between the original and new hosts to spread, as they can only survive outside the host from hours to days. In contrast, some enteric viruses can survive outside the host from weeks to months, allowing them to spread beyond the habitat of the reservoir host through fomites or environmental transport [159–161].

Cross-host exposures are an important step in zoonosis, and some spillover events are likely prevented because of limited contact between the pathogen and the potential new hosts due to geographical, ecological, and behavioural separation. However, the changes in human behaviours, such as interactions with wild animals and extensive animal and vegetation farming, may overcome these barriers and increase the risk of zoonotic spillovers. For example, wildlife trading of bats, pangolins, raccoon dogs and civet cats is thought to have played a critical role in the outbreak of both SARS-CoV-1 and 2 [162]. However, most zoonotic events in humans are from domesticated animals due to closer proximity. Zoonotic cases of Nipah virus among pig farmers in Malaysia in 1998 were hypothesized to be the result of intensive fruit farming, attracting fruit bats, the natural hosts, to come in contact with farmed pigs [163]. Human cases of avian influenza infections most often arise after initial outbreaks in poultry farms, where spillover events between wild birds and farmed birds occur [164, 165].

Climate change is believed to have increased zoonotic risk by altering the population and behaviour of reservoir hosts, increasing their chance of coming in contact with humans. For example, global warming was linked to an increased mosquito population, increasing the chance of exposure to vector-borne pathogens [166, 167]. The increase in rodent numbers in the mentioned hantavirus outbreak is attributed to a sudden increase in precipitation also linked to global warming [156]. Temperature has also been shown to affect the viability of viruses outside their host. For example, coronaviruses and influenza viruses tend to maintain viability at lower temperatures, which partially explains why the SARS-CoV-1 epidemic vanished in summer and flu season occurs in winter [168–170].

The role of Species-specific differences in spillover events

After a recipient host has been exposed, the species-specific barriers in the new host and their interactions with the pathogen will determine whether an infection will be established at a given dose. These barriers vary depending on the specific combinations of pathogens, species and individuals. To establish infection, a pathogen must have the ability to efficiently infect and replicate in the appropriate cell type or tissue in the new host, which can be restricted at many different levels. Physiological barriers such as differences in body temperature and pH may affect pathogen viability. Cellular barriers include the lack of receptors and other host dependency factors involved in binding and entry, cellular trafficking, genome replication, and gene expression. Differences in IFN-induced and other immune responses can also play a role in restricting cross-species infections. Although these barriers in the new hosts might not be able to completely block out zoonotic pathogens, they will nevertheless render them less fit or stable. These pathogens would therefore require one or more corresponding changes to adapt and

establish disease in the new host, making the species-specific barrier increasingly difficult to cross.

One may be tempted to assume that the chance for spillover events is higher between more closely related species. For instance, spillover events between two species of mammals are more common than events involving a mammal and a reptile. Indeed, HIV, one of the best-studied zoonotic pathogens in humans, has originated from simian immunodeficiency virus (SIV) which was circulating in non-human primates. Four HIV-1 groups were identified in humans and phylogenetic analysis suggests they each resulted from an independent cross-species transmission event [171]. The precursor of all four groups is SIVcpz in chimpanzees, which is the result of recombination events between three different SIVs, of which two have confirmed monkey origin. SIVcpz was then transmitted to gorillas and humans, giving rise to SIVgor, and HIV-1 groups M and N respectively. Two further zoonotic transmission events of SIVgor to humans resulted in HIV-1 groups O and P. Given the closer genetic relationship between humans and great apes (i.e. chimpanzees and gorillas) compared to monkeys, it is plausible that SIVcpz and SIVgor are more readily to overcome the species-specific barriers posed by humans than SIVs in monkeys. This suggests the greater apes acted as intermediate hosts for monkey SIV to infect humans [172]. None of the SIVs infecting numerous monkey species have been detected in humans, perhaps due to the greater species gap with one striking exception; SIVsmm from sooty mangabey monkeys. SIVsmm has been transmitted to humans by at least nine independent zoonotic events, giving rise to HIV-2 groups A to I. This exception is probably multi-factorial. Apart from these monkeys having a high prevalence for SIVsmm (i.e high pathogen pressure) and being popular

household pets and bushmeat sources (i.e. higher chance of human exposure), SIVsmm viral proteins appear to be more capable of overcoming human antiviral ISGs without the need of adaptation [172]. Nonetheless, the fact that HIV-1 group M is better at countering ISGs such as TRIM5α and tetherin, evading PRRs such as cGAS, as well as utilizing host dependency factors such as RanBP2, on top of higher pathogenicity and prevalence than other HIV-1 groups and HIV-2 in humans, suggests their success is due to better adaptation to the cross-species barriers posed by humans [172–174].

Nonetheless, spillover events between distantly related species occur much more often than we originally thought, especially for pathogens that have a wider host range in the first place. IAV is an example of this, with the virus infecting birds and many different mammals including humans [155]. Although IAV is more readily able to infect different species than other pathogens that have a narrower host range, there are still multiple species-specific barriers the virus has to overcome before a spillover event can occur. The first species barrier is the difference in the cell attachment factor sialic acid. For example, the haemagglutinin (HA) of avian IAV preferentially binds α2-3-linked sialic acid, meaning the virus is inefficient at infecting the upper respiratory tract (URT) of humans, where the majority of sialic acid is α-2-6-linked [175]. Similarly, non-human primates and mice which have higher expression of α2-3-linked sialic acid are poor animal models for human-adapted IAV [176]. After attachment, IAV is endocytosed and the drop in pH within the endosome triggers an irreversable change in HA which is critical for the release of viral RNA into the cytoplasm. The HA of human-adapted IAV is more stable at lower pHs, allowing the virus to maintain viability in the mildly acidic environment in the URT. In contrast, the HA of avian IAV is more sensitive to lower

pH, prompting premature conformational change before entry [177]. The temperature of the human URT is around 32°C, whereas the temperature of the enteric tract in birds is 40°C. This temperature difference renders the polymerase enzyme of avian IAV less active, resulting in ineffective virus replication [178, 179]. Apart from temperature, the polymerase activity of IAV is also influenced by host factors such as ANP32A. Mammalian ANP32A lacks a 33 amino acid motif found only in avian ANP32A which is crucial for interacting with the polymerase of avian IAV and enhancing its activity. Avian IAV therefore replicates poorly in mammalian cells without adapting to the shorter cellular factor [180]. The relatively large phylogenetic gap between birds and mammals explains not just the differences in certain cellular factors but also the IFN response. A recent study has identified an ISG in humans that can potently restrict IAVs of avian origin but not human. This antiviral ISG, BTN3A3 was shown to be one of the major barriers to spillover events in humans, as most avian IAVs that have successfully crossed the species barrier have developed resistance towards the gene [181].

As shown by the example of HIV and IAV, the IFN response and ISGs play a critical role as barriers against zoonotic events. Viral IFN evasion mechanisms often rely on direct protein-protein interactions. Species-specific differences can disrupt these interactions rendering the evasion mechanism ineffective, contributing to the barrier against zoonotic events. In addition, the ISG profiles of different species are divergent. A recent study showed that, among hundreds of ISGs that were upregulated following IFN response, only 62 were conserved in the ten vertebrate species, with an additional 28 in the nine mammalian species they have investigated [154]. This is probably the result of a continuous arms race between pathogens and their hosts. The primate-specific BTN3A3

discussed above is a perfect example. Intriguingly, the upstream IFN system in general is not identical among species, especially the type I IFNs. Besides the 5 subclasses of type I IFN mentioned in the previous section which is commonly found in most mammals, some IFN subclasses are more species-specific. For example, IFNδ is porcine and equine specific, IFNζ and IFNτ are murine and ruminant specific respectively, while IFNαω is found in porcine, equine and bovine species [182]. Moreover, some of these subclasses can be further diversified into multiple subtypes, and the number of subtypes in each species can be different. For instance, there is only one functional gene for IFNω in humans, while cattle, pigs, horses, and cats have multiple IFNω genes [183, 184]. It was originally assumed that these subclasses and subtypes have a similar function as they all bind to IFNAR1/2 and should trigger the same signalling cascade. However, there is increasing evidence they appear to have different antiviral potency in various cell types and tissues against different viruses. For example, a study has shown that some porcine IFNα subtypes are better at suppressing "contemporary " viruses such as porcine reproductive and respiratory syndrome virus (PRRSV) and vesicular stomatitis virus (VSV), while some subtypes are better at suppressing the expression of "ancient" porcine endogenous retrovirus -γ1 [182].

1.3.2 Zoonotic potential of coronavirus

Theoretically speaking, all pathogens have the potential to cross the species barrier. However, pathogens that can mutate and adapt at a faster pace have a higher chance of causing spillover events. Viruses, and in particular RNA viruses have been the major source of zoonotic outbreaks in the past few decades, probably due to their high mutation

rate, ranging from 10^{-3} to 10^{-6} nucleotide substitutions per site per year compared to 10^{-6} to 10^{-8} for DNA viruses [185]. Apart from mutations in individual nucleotides, segmented RNA viruses such as IAVs can also generate mutants by reassortment. When two or more IAVs infect the same host cell, their segments may be exchanged during virion packaging, generating different combinations of reassortant progeny [186]. In fact, pigs have been proposed to be a 'mixing vessel' for avian and human IAVs as they are susceptible to both, providing chances for avian IAVs to gain key mutations required to infect humans [187].

Recombination is another mechanism that facilitates RNA virus adaptation and the most widely accepted model of RNA recombination is 'copy choice recombination' [188]. Similar to reassortment, co-infection of two viruses in the same host cell is required. In this process, the RdRp of RNA viruses switches from its original RNA template to another during replication while remaining bound to the nascent nucleic acid chain, generating a hybrid RNA molecule [186]. This process seems to be much more frequent in positive-sense RNA viruses when compared to negative-sense RNA viruses [189]. Apart from retroviruses, such as HIV, which are thought to have one of the highest recombination rates, recombination has also been shown to be very frequent in coronaviruses, probably due to their unique gene expression strategy of discontinuous transcription [186]. As discussed in the previous section, this strategy relies entirely on the template-switching property of coronavirus RdRp, which is precisely what recombination requires.

Mutation at the nucleotide level for coronaviruses is relatively low (i.e. 3.76×10^{-6} substitutions per site per passage compared to 9.01×10^{-5} for IAV) as they are the

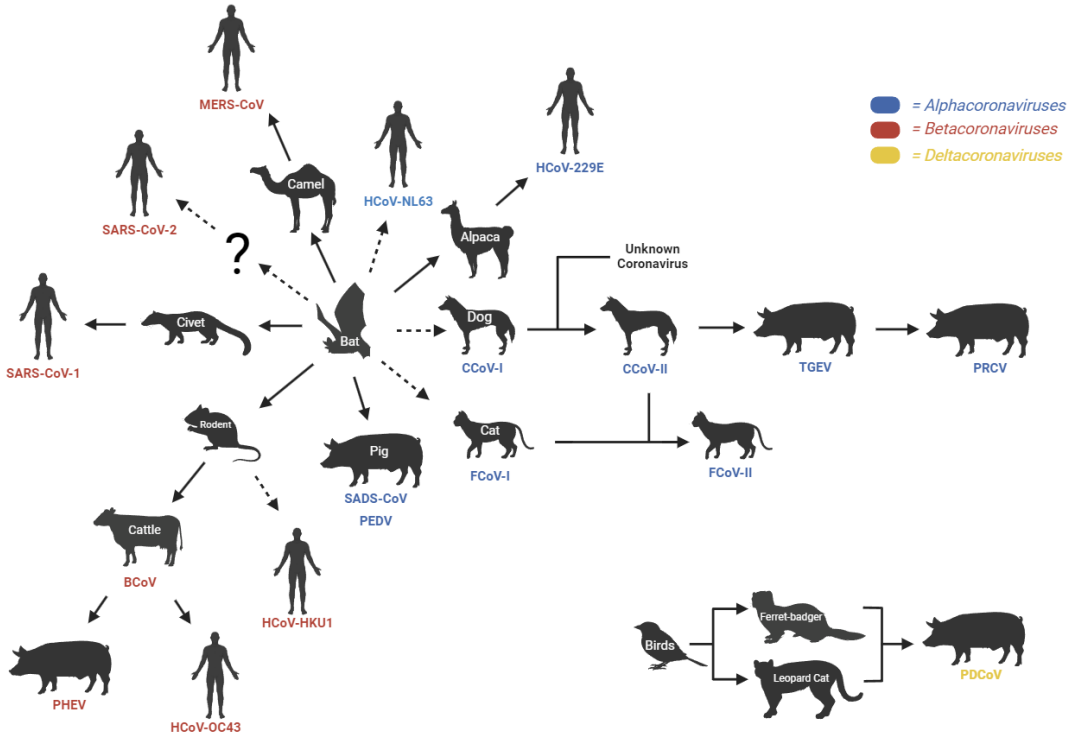


Figure 1.5: **Zoonotic transmission of human, swine, and related coronaviruses.** Bats are regarded as the original host of coronaviruses and therefore most coronavirus infections in other animals are believed to be a result of spillover events from bats. Humans and swine can receive coronavirus directly from bats or indirectly through an intermediate host. Porcine deltacoronavirus (PDCoV) is the only exception where birds are regarded as the origin host for deltacoronavirus. Solid arrows represent transmissions supported by direct evidence and dashed arrows represent hypothesised transmissions. Alpha-, beta-, and deltacoronavirus were represented by blue, red and yellow font respectively (Created with BioRender.com).

only family of RNA viruses that have proofreading machinery, probably through the exonuclease activity of nsp14 [23, 190]. Nonetheless, evidence of recombination within the genomes of coronaviruses is extremely common, as a single genome can have traces of multiple ancestral coronaviruses from different evolutionary sources. Recombination involving the spike gene in particular, is believed to be the driving force of cross-species events and disease emergence in novel host species, as receptor compatibility remains the biggest species barrier [191, 192]. Although the role of recombination in the emergence of SARS-CoV-2 is still under debate, there is compelling evidence of

spike recombination for the other 6 human coronaviruses. For example, SARS-CoV-1-like viruses found in bats cannot utilize ACE2 as entry receptors and have little to no capacity to infect humans. These viruses likely acquire the ability to use ACE2 through spike recombinations before causing an outbreak in humans [193, 194]. Similarly, MERS-CoV-like viruses cannot bind to DPP4 and may have gained this function by recombination [195]. Spike recombinations are also observed in HCoV-229E, OC43, HKU1 and NL63, but their direct role in causing human infections is yet to be demonstrated [196–198]. Changes in the spike gene can sometimes have implications beyond host specificity, as they are also involved in virulence. For example, the decreased virulence of the SARS-CoV-2 Omicron variant is attributed to mutations in the S gene [199].

Spike recombination has not just promoted animal-to-human transmission of coronavirus but has also expanded the host range of some veterinary coronavirus, which can lead to severe disease in companion and farm animals. The species *Alphacoronavirus 1* comprises the aforementioned porcine TGEV and PRCV, as well as feline coronavirus (FCoV), canine coronavirus (CCoV) and probably a poorly defined enteric coronavirus found in rabbits [6]. These viruses display greater than 96% sequence identity in the replicase gene but only 56% in the spike ectodomain, suggesting they share a common ancestor before recombination in the spike gene which leads to diseases in their respective animals [200, 201]. It is believed serotype I FCoV and CCoV share a common ancestor, and a spike recombination event with an unknown coronavirus led to the emergence of serotype II CCoV which has a distinct S protein. Further independent spike recombination events between serotype II CCoV and serotype I FCoV have led to the emergence of serotype II FCoV and recombinant serotype I/II CCoV. TGEV is believed to have

evolved from serotype II CCoV and further deletion has led to the emergence of PRCV (**Fig 1.5**) [202]. Recent spike recombination between a hypervirulent pantropic CCoV and FCoV has resulted in a sharp increase of feline infectious peritonitis cases in the cat population of Cyprus, leading to thousands of deaths [203]. TGEV has also been found to recombine with another alphacoronavirus, PEDV, leading to the emergence of an unnamed recombinant swine enteric coronavirus with distinct pathogenic properties [204].

Although bats are regarded as the original host of most, if not all, coronaviruses, direct bat-to-human transmission has not been reported to date [205], rather, intermediate hosts such as civet cats and dromedary camels were shown to play a critical role in the SARS-CoV-1 and MERS-CoV outbreaks respectively. The outbreak of SARS-CoV-2 almost certainly involved an intermediate host, with some phylogenetic studies speculating pangolin as a potential candidate [206]. Some HCoVs that cause milder infections were also found to have been transmitted to an intermediate host before becoming endemic in humans. For example, HCoV-OC43 is believed to have originated from bovine coronavirus, suggesting cattle-to-human spillover. Meanwhile, alpacas and rodents were suggested to be the intermediate hosts of HCoV-229E and HKU-1, respectively (**Fig 1.5**) [205, 207]. These examples suggest intermediate hosts can facilitate the spillover of coronaviruses from their original host (i.e. bats) to humans, likely because they **(1)** may serve as a mixing vessel to allow recombination with other coronaviruses and gain the ability to infect humans; **(2)** may act as a stepping stone by closing the species gap and enhancing pathogenicity towards human; or **(3)** have closer proximity to humans due to geographical and behavioural reasons.

1.4 Aims

In previous sections, the interplay between viruses and the innate immune response was discussed in detail. However, these are almost entirely based on studies carried out in humans or mice. Despite some aspects of the porcine immune system resembling that of humans, some fundamental differences still exist [208]. For instance, key components of the inflammasome pathways in humans such as AIM2 and NLRC4 are missing in pigs [209]. These differences can have important implications, as AIM2 was reported to be a critical DNA sensor for the efficacy of DNA vaccines against IAV [210]. Although the structural motifs of other immune-related genes such as TLRs are relatively well conserved in pigs, functional studies of the porcine PRR repertoire involved in responses to PAMPs and other ligands are lacking. These cannot be overlooked as PRRs can have altered receptor specificity despite sequence conservation. For example, imiquimod, an immune response modifier activating TLR7 in humans, will also activate TLR8 in pigs [211].

Similarly, studies dedicated to understanding porcine antiviral ISGs are relatively lacking and are mostly based on existing knowledge of their human or mouse counterparts. Porcine ISGs that have been studied in regard to their direct antiviral potency include ISG15, IFITM1, MX1, GBP1 and more [212–215]. All of these have been studied in detail in humans. While these studies contributed to the understanding of functional similarities and differences between homologous antiviral genes, they cannot address the intrinsic differences between the porcine and human repertoire of antiviral genes. Such studies are lacking in the field and this project attempts to fill in some of these knowledge

gaps by interrogating the antiviral potential of a wide range of porcine ISG, including candidates with both known and unknown antiviral potency in other species.

The aim of this project is to understand how the porcine IFN response restricts coronavirus infection and replication. Beyond this, we aim to use comparative ISG screens to identify porcine ISGs that inhibit SARS-CoV-2, a virus that has no evolutionary history of infecting pigs and would therefore be predicted to be less adapted to replication in porcine cells, but not the porcine-adapted coronavirus PRCV. We hypothesise that such ISGs would contribute to the evolutionary barrier against viruses jumping from other species into pigs.

The IFN response is pleiotropic in nature, meaning that the majority of antiviral effects are contributed by more than one ISG or antiviral factor. Loss-of-function studies such as the use of small interfering RNAs (siRNA) and microRNAs or the more recent clustered regularly interspaced short palindromic repeats (CRISPR) -Cas9 gene editing technology alone may not be effective in dissecting the antiviral potential of individual ISG, as the antiviral effects are likely to be redundant. Gain-of-function ectopic expression of ISGs to mimic IFN-mediated induction has therefore become a complementary way to tackle this problem. Early attempts applied a small number of known ISGs to test for their specific antiviral properties against individual viruses [216, 217]. In 2011, Schoggins *et al.* compiled a list of 389 ISGs and screened them against a wide range of viruses, successfully identifying a diverse range of novel effectors of the type I IFN response [109]. Since then, the use of such ISG libraries has led to a much better understanding of the IFN response.

The One Health concept suggests human health is closely related to animal and environmental health. Investigating virus-host interaction and IFN response in humans and other animals is equally important to understanding zoonotic diseases. Following the success of large-scale screens in identifying antiviral factors using a library of human ISGs against a diverse panel of viruses, we sought to expand the scope of these studies to other animal species. Pigs are important livestock species that are genetically, anatomically, physiologically, and immunologically similar to humans and have been used as animal models for different research in addition to mouse and non-human primates [218]. Moreover, they are hosts for viruses with high zoonotic risk such as IAV, Nipha virus and coronaviruses, making them potential mixing vessels for future pandemics. The ability of SADS-CoV to infect and replicate in primary human cells and human cases of PDCoV infection in Haiti further highlighted the spillover risk of newly emerged coronaviruses in pigs [219, 220]. We have therefore developed an arrayed porcine ISG expression library containing 432 individually cloned type I ISGs on the same lentivirus backbone used for the human library based on published RNA-seq and screening data [154, 221].

As mentioned in the previous sections, viruses need to overcome barriers posed by the evolutionary difference between species, such as the IFN response, to establish infection in another species. In this project, we will apply the porcine ISG library to PRCV, an endemic porcine coronavirus that is highly adapted to pigs, and SARS-CoV-2, a human coronavirus that can replicate efficiently in some porcine cell lines *in vitro* but fails to infect pigs *in vivo* [222]. Pigs infected with PRCV have been used as a model to study the effect of corticosteroid treatment in SARS-CoV-1 patients in the past due to the

similarity in pathology, cell tropism, and clinical outcomes of the two viruses in their respective hosts [223, 224]. We hypothesize the endemic PRCV is more adapted and better at counteracting the porcine IFN response than the non-adapted SARS-CoV-2. By comparing the two viruses, we aim to identify novel porcine-specific antiviral factors and potentially anti-coronavirus factors that contribute towards barriers against zoonotic transmission of coronavirus.

CHAPTER 2

Materials and methods

2.1 Materials

2.1.1 General

PrimeSTAR GXL DNA polymerase	Takara Bio, R050A
SalI-HF	New England Biolabs, R3138
NotI-HF	New England Biolabs, R3189
XhoI	New England Biolabs, R0146
T4 DNA ligase	New England Biolabs, M0202
QIApquick PCR purification Kit	QIAGEN, 28104
PureLink Quick Gel Extraction Kit	Inviorgen, K220001
PureLink HiPure Plasmid Maxiprep Kit	Invitrogen, K210007
PageRuler prestained protein ladder	Invitrogen, 26616
CellTiter-Glo Luminescent Cell Viability Assay	Promega, G7570
Triton X-100	Sigma-Aldrich, X100
WGA with Alexa Fluor 633 Conjugate	Invitrogen, W21404
DAPI Solution	Invitrogen, 62248
Lenti-X™ Provirus Quantitation Kit	Takara, 631239
COVID-19 Rapid Antigen Nasal Lateral Flow Test Kit	FlowFlex, LO31-118M5

2.1.2 Buffers and Solutions

4X FA Rinse Buffer	107mM Na_2CO_3, 400mM $NaHCO_3$, 145mM NaCl, pH9.0
4X Protein loading buffer	40% Glycerol, 240mM Tris-HCl pH6.8, 8% SDS, 0.04% bromophenol blue, 5% β-mercaptoethanol
5X siRNA Buffer	Dharmacon, B-002000-UB-100
10X Western SDS Running Buffer	250mM Tris Base, 1920mM Glycine, 1% SDS, pH8.3
10X Western Transfer Buffer	250mM Tris Base, 1920mM Glycine, 20% Methanol
10X Tris-Buffer Saline (TBS)	198mM Tris Base, 1.5M NaCl, pH7.6
TBS-Tween (TBS-T)	1% Tween20 in TBS
Phosphate-Buffered Saline (PBS)	138mM NaCl, 2.7mM KCl, 8mM Na_2HPO_4, 1.5mM KH_2HPO_4, pH7.6
Virus Lysis Buffer (VL buffer)	2.5% IGEPAL CA-630, 150mM NaCl, 1:4000 RNasin Plus, in 10mM Tris-Cl, pH7.5
Western Blocking Solution	5% Milk in TBS-T
Mounting Fluid	Glycerol/FA Rinse Buffer, 1;1

2.1.3 Cell Culturing

Dulbecceo's Modified Eagles's Medium (DMEM)	Sigma-Aldrich, D5796
DMEM/F12	Gibco, 11320033
RPMI-1640	Sigma-Aldrich, R5886
Fetal Bovine Serum (FBS)	Gibco, 192-1005PJ
Horse Serum	Sigma-Aldrich, H1138
Reduced Serum Medium (Opti-MEM)	Gibco, 11058021
Penicillin-Streptomycin (10,000 U/mL; Pen-Strep)	Gibco, 15140-122
Antibiotic/Antimycotic Solution (100x)	Capricorn Scientific, AAS-B
Gentamicin (10 mg/mL)	Gibco, 15710064
Collagenase type I	Sigma-Aldrich, SCR103
Trypsin-EDTA	Sigma-Aldrich, T3924
DEAE-Dextran	Supelco, 93556
Lipofectamine 2000	Invitrogen, 11668019
Lipofectamine RNAiMax	Invitrogen, 13778150
Lenti-X Concentrator	Takara Bio, 631231
Porcine Interferon Alpha	PBL Assay Science, 17105
Universal Type I Interferon	PBL Assay Science, 11200

2.1.4 Cell lines and Viruses

Cell lines

ST —Swine testis cells, a gift from Dr Christine Tait-Burkard (Roslin Institute, University of Edinburgh), purchased from ECACC

NSK —Newborn swine kidney cells, a gift from Dr Simon Lillico (Roslin Institute, University of Edinburgh)

NPTr —Newborn pig trachea cells, a gift from Dr Simon Lillico

PK15 —Porcine Kidney cells, a gift from Dr Simon Lillico

NT2 —Epithelial-like human testis carcinoma, a gift from Professor Rod Mitchell (Institute for Regeneration and Repair, University of Edinburgh)

TM4 —Mouse Sertoli cells, a gift from Professor Rod Mitchell

Vero —African green monkey kidney epithelial cells a gift from Professor Jürgen Haas (Institute for Regeneration and Repair, University of Edinburgh)

Lenti-X 293T —A subclone of human embryonic kidney cells expressing mutated SV40 large T antigen optimised for lentivirus production, purchased from Takara Bio

Caki-1 —Human kidney clear cell carcinoma, a gift from Dr Christine Tait-Burkard, purchased from AMSBio

Calu-1 —Human lung epidermoid carcinoma, a gift from Dr Christine Tait-Burkard, purchased from ECACC

BEAS-2B —Human bronchial epithelial, a gift from Dr Christine Tait-Burkard

MRC-5 —Human lung fibroblast, a gift from Professor Jürgen Haas

NHDF —Normal human dermal fibroblast, purchased from Gibco

Viruses

PRCV strain ISU-1 —a gift from Dr Toby Tuthill (Pribright Institute)

PRCV strain AR310 —a gift from Dr Christine Tait-Burkard, purchased from ATCC

TGEV strain Miller (TC-adapted) —a gift from Dr Christine Tait-Burkard, purchased from ATCC

SARS-CoV-2 isolate EDB2 —Lineage B.1, early isolate, a gift from Dr Christine Tait-Burkard

SARS-CoV-2 Delta variant —Lineage B.1.617.2, a gift from Dr Christine Tait-Burkard

SARS-CoV-2 Omicron variant —Lineage BA.1, a gift from Dr Christine Tait-Burkard

2.1.5 RNA extraction and RT-qPCR

TRIzol Reagent	Invitrogen, 15596026
TRIzol-LS Reagent	Invitrogen, 10296028
TURBO DNA-Free kit	Invitrogen, AM1907
High-Capacity cDNA Reverse Transcription Kit	Invitrogen, 4368814
GoTaq 1-Step RT-qPCR System	Promega, A6020
2X TaqMan qPCR Master Mix	Invitrogen, 4444557
SensiFAST SYBR Hi-ROX kit	Bioline, BIO-92005
TaqMan Gene expression assays	Invitrogen, see below

Target	Gene	Assay ID
Human	ACE2	Hs01085333_m1
Human	GAPDH	Hs02786624_g1
Human	ISG20	Hs00158122_m1
Human	ZCWPW1	Hs01553931_m1
Mouse	GAPDH	Mm99999915_g1
Mouse	ISG20	Mm00469585_m1
Mouse	ZCWPW1	Mm01196994_m1
Pig	ACE2	Ss03390186_m1
Pig	APN	Ss03394557_m1
Pig	GAPDH	Ss03375629_u1
Pig	ISG20	Ss06942055_g1
Pig	ZCWPW1	Ss06898521_g1

2.1.6 Oligo Sequences

Name	Sequence 5'-3'
Virus Detection	
PRCV_N_F	TGA AGG GCC AAC GTA AAG AG
PRCV_N_R	CAA CCC AGA CAA CTC CAT CTA A
PRCV_N_Probe	ACT TAG GCA CTG GAC CTC ATG CAG
PRCV_ORF3_F	GAG CTG CAC CGT TTA AGA GA
PRCV_ORF3_R	GCA TGA GTT AAG CCA CGT ATT G
PRCV_S_F	GCC ATT GAT TTA TGG AGA
PRCV_S_R	GTA TAA AAC CTC CTG GCT GT
SARS-CoV-2_N3_F	GGG AGC CTT GAA TAC ACC AAA A
SARS-CoV-2_N3_R	TGT AGC ACG ATT GCA GCA TTG
pAPN Cloning	
Xba1_pAPN_F	AGT CTA GAA TGG CCA AGG GAT TCT ACA TTT CCA A
BamH1_pAPN_R	GAC GGA TCC TTA GCT GTG CTC TAT GAA CCA ATT CAA C
pLenti_HA_pAPN_R	GAC GGA TCC AGC ATA ATC TGG AAC ATC ATA TGG ATA GCT GTG CTC TAT GAA CCA ATT
siRNA	
ssZCWPW1_#1	AAG AGG CAG AGC AGA UCA GUA & AAU ACU GAU CUG CUC UGC CUC

ssZCWPW1_#2	AAU UAG ACA AGG AAC CAG GAA & AAU UCC UGG UUC CUU GUC UAA
ssZCWPW1_#3	AAG GAA GAG GAA GAG UCA GAG & AAC UCU GAC UCU UCC UCU UCC
siNeg	ACG TGA CAC GTT CGG ACA ATT & UUC UCC GAA CGU GUC ACG UTT
pSCRPSY Sequencing	
pSCRPSY_F	TGG AAC TTC TGG GAC GCA GG
pSCRPSY_R	TGG TAG ATC CAC AGA TCA AGG

2.1.7 Antibodies

Name	Supplier and Catalog No	Dilution
FITC TGEV polyclonal antiserum	VMRD, CJ-F-TGE-10ML	1:1
Mouse anti-TGEV Coronavirus Spike glycoprotein [1AF10]	Absolute Antibody, Ab02090	1:1000
Mouse Anti-Porcine Coronavirus Spike Antibody (5B4)	Native Antigen, MAB12431	1:1000
Mouse Coronavirus pan Monoclonal Antibody (FIPV3-70)	Invitrogen, MA1-82189	1:1000
Mouse HA Tag Monoclonal Antibody (2-2.2.14)	Invitrogen, 26183	1:1000
Mouse SARS/SARS-CoV-2 Nucleocapsid Monoclonal Antibody (B46F)	Invitrogen, MA1-7404	1:100
Rabbit Recombinant Anti-CD13 antibody [EPR4058]	Abcam, ab108310	1:1000
Rabbit Recombinant Anti-ACE2 antibody [EPR4436]	Abcam, ab108209	1:150
Rabbit GAPDH Polyclonal antibody	Proteintech, 10494-1-AP	1:5000
Rabbit ZCWPW1 Polyclonal Antibody	Invitrogen, PA5-96731	1:1000
Goat IRDye 800CW anti-Mouse	Li-cor, 925-32210	1:10000
Goat IrDye 680RD anti-Rabbit	Li-cor, 926-68071	1:10000

2.2 Methods

2.2.1 Cell Culturing

Maintenance

Caki-1 cells were maintained in RPMI-160 medium supplemented with 10% FBS and 1% Pen Strep. TM4 cells were maintained in DMEM/F12 supplemented with 5% horse serum and 2.5% FBS. All other cells were maintained in DMEM supplemented with 10% FBS and 1% Pen Strep (complete medium) unless otherwise specified. FBS and horse serum were heat-inactivated at 56°C for 30 minutes before use. Cells were maintained at 37°C in 5% CO_2 and regularly passaged every 3 to 4 days to avoid over-confluence.

Freezing and resuscitation

Cells were washed with PBS, trypsinised and resuspended in complete medium. Cell pellets were collected by centrifugation at 1200 x *g* for 5 mins. Supernatants were discarded and cell pellets were re-suspended in 20% DMSO in FBS, aliquoted in cryo-tubes and stored in a freezing container at -80°C for slow freezing before being moved to liquid nitrogen storage.

To resuscitate cells, medium pre-warmed at 37°C was added into cryo-tubes. The contents in the cryo-tubes were pipetted up and down to facilitate thawing before being transferred into a cell-culturing flask.

siRNA Transfection

RNA oligos were ordered from Integrated DNA Technologies (IDT). The passenger and guide strands were resuspended into 100 μM and mixed with 5X siRNA buffer (Dharmacon) and nuclease-free water at a ratio of 1:1:1:2 to generate 20 μM of siRNA complex. The mixture was heated at 95°C for 2 minutes on a heat block and gradually cooled down by turning off and leaving it on the heat block until it reached 40°C. The annealed siRNA complex was further cooled down at room temperature for 30 minutes before being aliquoted and stored at -20°C until use.

Cells were seeded at the same time as transfection. For each well of a 24-well plate, 1ul of 20μM siRNA diluted in 50μl Opti-MEM was mixed with 1μl of Lipofectamine RNAiMAX diluted in another 50μl of Opti-MEM. The mixture was incubated at room temperature for 5 minutes and reverse-transfected into the well. A second transfection was done 24 hours after the first transfection as described. The states of the cells were accessed 3 days after the second transfection.

To assess the viability of transfected cells, transfection was done in 96-well plates similarly with a dilution factor of 0.25. Cell viability was measured by CellTiter-Glo assay (Promega) following the manufacturer's protocol.

Lentivirus production and transduction

15 μg of transfer plasimd, 12μg of psPAX2 (Addgene #12260) and 3μg of pMD2.G (Addgene #12259) were diluted in 1.8mL of Opti-MEM while 75μl of Lipofectamine 2000 were diluted in 1.8mL of Opti-MEM separately. After 5 minutes of incubation

at room temperature, the two components were mixed and incubated for a further 20 minutes at room temperature. The transfection mix was then added to a T75 flask of 80% confluence HEK 293T cells. Supernatants were harvested 48 and 96 hours post-transfection with fresh medium replaced in between. The harvested supernatant was spun to remove cell debris before passing through a 0.45μm filter. The resulting lentiviruses were concentrated by Lenti-X concentrator (Takara Bio) following the manufacturer's protocol. Concentrated lentiviruses will be aliquoted and stored at -80°C until use.

To transduce cells, lentivirus was thawed at room temperature and diluted by a variable volume of complete medium based on titration result. The diluted lentivirus was then supplemented by 16μg/mL of DEAE-Dextran before being added to cells at >90% confluency. For transduction in 96-well plates, the plates were spun at 800 x *g* for 10 minutes to enhance efficiency. The medium of transduced cells was changed after 24 hours to reduce the chance of cytotoxicity. Transduction levels were accessed by Cytation 3 (BioTek) at a gain of 120 for RFP and 66 for GFP. The state of cells was accessed 48 to 72 hours post-transduction.

2.2.2 Virus Work

Propagation

For PRCV, ST cells were grown in T175 for 3 to 4 days to ensure 100% confluency. The virus stock was diluted in 10ml of serum-free medium (SFM; DMEM without FBS supplement) pre-warmed at 37°C and used to infect the cells at a multiplicity of infection (MOI) of 0.01. After 2 hours of incubation at 37°C on a rocker, the inoculum was

removed and cells were washed with PBS once before 20ml of complete medium was added. The cells were examined daily until complete cytopathic effects (CPE) were observed (3 to 4 days). The supernatant was harvested and spun at 1200 x *g* to remove any cell debris before being aliquoted and stored at -80°C until use.

For SARS-CoV-2 EDB2, passage 2 virus was generated using either Veros (for all *in vitro* infections) or hACE2-NSK cells (for *in vivo* infections). The details of virus isolation from CoVID-19 patients and the first 2 passages (P0 and P1) have been described previously [225]. Veros or hACE-2 NSK cells were grown in T75 for 2 to 3 days to ensure 100% confluency. The virus stock was diluted in 5ml of SFM pre-warmed at 37°C and used to infect cells at a multiplicity of infection (MOI) of 0.01. After 2 hours of incubation at 37°C on a rocker, the inoculum was removed and cells were washed with PBS once before 10ml of complete medium was added. The supernatant was harvested 2 days post-infection (DPI) and spun at 1200 x *g* to remove any cell debris before being aliquoted and stored at -80°C until use.

Infection

PRCV or SARS-CoV-2 stocks were diluted in SFM pre-warmed at 37°C to desired concentration. Cells were washed 2 times with room temperature PBS before the addition of inoculum to remove residue serum. A minimal amount of inoculum (50 µl for 96-well plates and 200 µl for 24-well plates) was added to maximize virus attachment efficiency and was incubated at 37°C for 2 hours on a rocker. Fresh complete medium was added after removing the virus inoculum.

Quantification by Two-step RT-qPCR

All primers for PRCV detection were designed based on the genome of PRCV strain ISU-1 (Genbank accession DQ811787). The cDNA samples were prepared as described below (see 2.2.3) and were used for qPCR reaction using the SensiFAST SYBR Hi-ROX kit following the manufacturer's protocol to assess the efficiency of the designed primers. For the determination of PRCV nucleocapsid gene copy number in samples, a probe was further designed based on the above primers targeting N (see materials), and the custom TaqMan assay was purchased from IDT. The expected amplicon (111bp; nucleotide 26116-26226) was purchased from IDT and cloned into a pCDNA3.1 plasmid. RNA extraction and two-step RT-qPCR were performed as described below (see 2.2.3). A standard curve was generated every time using the plasmid control and the copy numbers were calculated using the equation as follows:

$$\text{Result (copies/ml)} = \frac{\text{Copies/µl x Intial Elution Volume (µl)}}{\text{Sample Volume (ml)}}$$

Quantification by direct lysis one-step RT-qPCR

The detailed protocol has been described previously [226]. Briefly, the supernatant from infected cells was lysed by an equal volume of VL buffer (see materials) by pipetting thoroughly up and down. The lysate was diluted 10X in nuclease-free water after 20 minutes of lysing at room temperature. 2 µl of the diluted lysate was added to 8 µl of RT-qPCR reaction prepared using the GoTaq 1-step RT-qPCR System (Promega) following the manufacturer's protocol containing 350nM of each, forward and reverse

primers targeting the N gene of PRCV or SARS-CoV-2 (see materials). The reaction was run according to the manufacturer's instructions on Rotor-Gene Q (Qiagen) using an annealing temperature of 60°C.

Quantifiaction by TCID50

ST (for PRCV titration) or hACE2-NSK (for SARS-CoV-2 titration) cells were seeded in 96-well plates 2 to 3 days in advance to ensure 100% confluency. A serial 10X dilution of virus samples was prepared using SFM pre-warmed at 37°C and 50µl of each dilution was added into 8 replicate wells. Following incubation at 37°C for 2 hours on a rocker, 50 µl of complete medium was added to the wells without the removal of virus inoculum. At 5 DPI, cells were fixed in 4% paraformaldehyde (PFA) for at least 20 minutes, before staining with 0.1% toluidine blue. The number of wells exhibiting CPE for each dilution was counted and the 50% tissue culture infectious dose (TCID50) was calculated using the Spearman–Kärber method.

Detection by antibody

Cells were seeded in 24-well plates with coverslips and infected by PRCV as described above. At 24 hours post-infection, coverslips were washed with PBS and fixed in 4% PFA for 20 min. After removing PFA, coverslips were washed twice before permeabilization by 0.1% Triton X-100 for 10 minutes. A wash was applied before staining the coverslips with fluorescein isothiocyanate (FITC) conjugated anti-TGEV polyclonal antiserum (VMRD) and incubated in a humid chamber at 37°C for 30 minutes. Coverslips were rinsed once and soaked in FA buffer (see materials) for 10 minutes. Wheat germ agglutinin with Alexa Fluor 633 Conjugate (Invitrogen) and DAPI in PBS were used as co-stain for

1 hour at room temperature. Coverslips were washed once before being mounted on glass slides with FA mounting fluid. Images were taken by Carl Zeiss LSM 710 confocal microscope and processed by Zen Blue (Carl Zeiss).

2.2.3 Molecular Biology

RNA extraction

The following volumes were optimized for samples from 24-well plates. For cell lysate samples, cells were washed with PBS once before being harvested by 500 µl of TRIzol; for supernatant samples, 250 µl of supernatant were lysed in 750 µl of TRIzol-LS. RNA extraction was then performed following the manufacturer's protocol. RNA was eluted in 45 µl of nuclease-free water and DNase-treated using the TURBO DNA-free kit. The concentration of samples was then determined by a NanoDrop spectrophotometer. All RNA samples were prepared from 24-well plates unless otherwise specified.

Two-step RT-qPCR

Approximately 1 µg of RNA was added into a 20 µl reaction of reverse transcription using the High-capacity cDNA reverse transcription kit (Invitrogen). 2 µl of cDNA from each reaction was added into a 10 µl qPCR reaction using 2X TaqMan qPCR Master Mix (Invitrogen) following manufacturer's protocol with appropriate TaqMan Primer Probe Assays. The reactions were run on Rotor-Gene Q (Qiagen) with conditions as follows: 50°C for 2 minutes; 95°C for 10 minutes; and cycle 95°C for 15 seconds and 60°C for 60 seconds 40 times.

Western Blotting

Cells were lysed and harvested in RIPA buffer and sonicated for 5 to 7 seconds. The samples were centrifuged at 15000 x *g* for 20 minutes at 4°C. The supernatant was harvested and the protein concentrations were determined by BCA assays. The samples were then mixed with 4X protein loading buffer and heated at 95°C for 5 minutes for denaturation before loading in a 10% SDS-PAGE gel. Protein was separated at 100V for 90 minutes and transferred to nitrocellulose membranes by wet transfer at 400mA for 120 minutes. The membrane was then blocked with 5% Milk in TBS-T for 30 minutes and probed with primary antibodies overnight at 4°C with rocking. The membranes were washed 3 times with TBS-T before secondary antibodies were applied for 1 hour. The blots were imaged by the Odyssey Fc imaging system (Li-Cor) after another 3 washes with TBS-T.

2.2.4 Generation of pAPN-NSK and hACE2-NSK cells

For the generation of pAPN-NSK cells, RNA and cDNA were prepared from NPTr cells using TRIzol as described above. Forward and reverse primers were designed based on the coding region of porcine aminopeptidase N (pAPN; GenBank accession KX342854) for the addition of Xba1 and BamH1 restriction sites respectively. The cDNA of pAPN was amplified using the PrimeSTAR GXL DNA polymerase (Takara Bio) following the manufacturer's protocol on the Veriti thermal cycler 96-well (Applied Biosystem) with conditions as follows: 98°C for 5 minutes; cycle 98°C for 15 seconds, 55°C for 30 seconds and 68°C for 3 minutes 35 times; and 4°C to stop the reaction. After confirming the size of the PCR product on an agarose gel, a second PCR was performed using the same

forward primer and a different reverse primer (pLenti_HA_pAPN_R; see material) to remove the last three nucleotides (i.e. the stop codon; to accommodate the P2A system in the target plasmid) and add an HA tag (YPYDVPDYA) at the carboxy terminus with the same conditions. The pLenti-CMV-Blast plasmid vector (Addgene #17486) and the PCR products purified by the QIApquick PCR purification kit (QIAGEN) were then digested by Xba1 and BamH1. The digested products were then run on an agarose gel and the desired bands were excised and purified by the PureLink Quick Gel Extraction Kit (Invitrogen). The insert was then ligated to the vector by T4 DNA ligase at a molar ratio of 3:1. The ligation product was transformed into competent *E.coli* bacteria (Stbl3) by 30 minutes incubation on ice, followed by heat shock at 37°for 3 minutes and another 2 minutes on ice. The bacteria were then grown on an LB agar plate with carbenicillin overnight at 37°C. Single colonies were picked and diluted in 50µl of distilled water and 3 µl of diluent was used as input for colony PCR. Colony with the expected insert size was grown in 100ml of liquid LB with carbenicillin overnight and the plasmids were purified using the PureLink HiPure Plasmid Maxiprep Kit (Invitrogen). Lentivirus was generated with the sequence-verified plasmid and transduced into Newborn Swine Kidney (NSK) cells as described above. 5 µg/ml of blasticidin was added to the cells for selection at 3 days post-transduction. Overexpression of pAPN in cells was confirmed by western blotting for the HA tag.

The hACE-NSK cells were generated with a similar approach. The plasmid encoding human angiotensin-converting enzyme 2 (hACE2) was a gift from Dr Christine Tait-Burkard. Briefly, the coding region of hACE2 was directly synthesized (Life Technologies) and cloned downstream of the CMV promoter, replacing the GFP in pLenti-CMV-GFP-

Hygro (Addgene #17446). 500 µg/ml of Hygromycin B was used to select transduced NSK cells.

2.2.5 Arrayed Lentiviral ISG screening

Library production

The cDNA library of 432 porcine ISGs cloned in pSCRSPY plasmids (GenBank accession KT368137) was prepared and sequence verified by Genwiz, Germany. The lentivirus library was generated in 96-well plate format. HEK 293T cells were seeded in eight 96-well plates (120 µl/well) to reach around 80% confluency the following day. Wells on the edges were not used to avoid evaporation (i.e. only 60 wells will be used per plate). For each well, 75ng of pSCRPSY, 75ng of psPAX2, 15ng of pMD2.G and 0.5 µl of lipofectamine 2000 were used for transfections. An additional pSCRPSY encoding GFP (pSCRPSY-GFP) was added to the cDNA library as a control (i.e. 433 reactions in total). The plasmids and lipofectamine were diluted in 30 µl of Opti-MEM separately. After 5 minutes of incubation, the two components were mixed thoroughly by pipetting up and down and incubated for 20 minutes further. The reaction mix (60 µl) was then added to the wells drop-wise to minimize disruption to the cell monolayer. The plates were then returned to the incubator for 48 hours.

From 2 to 5 days post-transfection, 100 µl of supernatant from each well was harvested into collection plates and 110 µl of complete medium was added back. The collection plates were stored at 4°C throughout the collection period. On day 6, all remaining supernatants were harvested and pooled with previous collections, resulting in a total of

~600ul per well. Lentivirus from pSCRPSY-EMP and GFP were used for titration to determine yield from the batch. The lentivirus library was then mixed thoroughly before being aliquoted into 4 sets (150 µl each), and stored at -80°C until use.

Library Screening

Unlike lentivirus produced from T75, the titer from 96-well plates is lower and cannot be concentrated. pAPN-NSK or hACE2-NSK cells were seeded at a lower density (~ 70% confluency) for transduction. The library was thawed at room temperature. After removing the medium, 30 µl of complete medium supplemented with 80 µg/ml DAE-Dextran (5X) was added into each well before 120 µl of the library was added and mixed gently. The plates were spun at 800 x *g* for 10 minutes before being returned to the incubator. The medium of transduced cells was changed 24 hours post-transduction. The cells were infected by PRCV or SARS-CoV-2 EDB2 at 3 days post-transduction. Virus titer was quantified at 1 DPI by direct lysis one-step RT-qPCR as described above on the Mx3000P (Agilent) system.

2.2.6 Challenge study of hACE2 transgenic pig

Generation of transgenic pig

pSL10-K18-hACE2 was used to generate the transgenic pigs. pSL10 is a modification of pLenti6/V5 D-TOPO (Invitrogen), with the ClaI-KpnI fragment removed and replaced with a cPPT/SV40 immediate early promoter / GFP / OPRE cassette. K18-hACE2 was PCR amplified from pK18-hACE2 (Addgene #149449)[227]. The amplified cassette was subcloned into BamHI and SalI sites in pSL10 to create pSL10-K18-hACE2. Lentivirus

was generated as described above.

Transgenic pigs were generated and genotyped as described previously [228]. Briefly, oocytes were collected by flushing the oviducts of super-ovulated Large-White gilts. 50 pl of lentivirus suspension was injected into the perivitelline space, which was then surgically implanted into 5 surrogate guilts. They were then mated with a wild-type breeder to generate F0 founder piglets.

Isolation of primary cells from transgenic pig

Ear notches biopsy with sizes of approximately 1 cm^3 from each transgenic pig was collected. Samples were washed in PBS and sliced into < 1 mm^3 sized pieces using scalpel blades. The tissue was cultured in digestion medium (DMEM supplemented with 20% FBS, 1X antibiotic/antimycotic, 50 µg/mL gentamicin and 0.5 g/ml collagenase I) in a T25 flask for 24 hours at 37°C with 5% CO_2. The sample was then filtered using a 70 µm cell strainer and cultured with outgrowth medium (digestion medium without collagenase I) at 37°C, 5% CO_2 until cells reached confluency.

In vivo SARS-CoV-2 infection study

Nine hACE2 transgenic and three wild-type 8-week-old pigs were transferred to Biosafety Level 3 (BSL3) housing. Following 7 days of acclimatisation, a blood sample was taken by jugular venepuncture and SARS-CoV-2 (2 ml of a 5 x 10^5 TCID50/ml suspension; isolate EDB2) was administered intranasally to each pig using a mucosal atomisation device (MAD300). Clinical scores and temperatures were recorded twice a day. At 2, 4 and 7 DPI, groups of 4 pigs (3 transgenic and 1 wild type) were euthanized by overdose

of sodium pentobarbital. A blood sample was taken by cardiac puncture and 2 nasal swabs were taken post-mortem. The first swab was tested immediately by lateral flow tests, and the second was stored in Trizol for later analysis. Tissue samples were taken at necropsy and stored immediately in 10% buffered formalin, lysed by Trizol, or stored in serum-free DMEM.

PCR

Gene expression was quantified by a two-step method. RNA was extracted and cDNA was prepared as described above (2.2.3). RT-qPCR reactions were performed using TaqMan Fast Advanced Master Mix (Applied Biosystems) and analysed with the Rotor-Gene Q real-time cycler (Qiagen). All gene expression values were normalized to the geometric mean of GAPDH, ACTB and RPL4 and are quantified by the $2^{-\Delta\Delta\ Ct}$ method unless otherwise specified. Assay ID for TaqMan assay (Applied Biosystems): Hs0108533_m1 (hACE2); Ss03375629_u1 (pGAPDH); Ss03376563_uH (pACTB); Ss03374067_g1 (pRPL4). SARS-CoV-2 N copy numbers were determined using the SARS-CoV-2 N1+N2 Assay kit (Qiagen, #222015) and standard curve generated from a positive control (Integrated DNA Technologies, #10006625).

ELISA assay

All sera were subjected to detection of antibodies against SARS-CoV-2 using ID Screen® SARS-CoV-2 Double Antigen Multi-species (Innovative Diagnostics) according to the manufacturer's protocol. The test detects antibodies to the nucleocapsid protein of SARS-CoV-2 for multiple species (i.e. minks, ferrets, cats, dogs, cattle, sheep, goats, horses and all other receptive species) with a specificity range of 97.8% to 100% as

reported by the manufacturer. The assay was validated when the optical density of positive control (OD_{PC}) was 0.35 and at least three times higher than the negative control (OD_{NC}). The optical density of each sample (OD_N) was used to calculate the S/P ratio value (expressed as %) where S/P = 100*(OD_N - OD_{NC})/ (OD_{PC} - OD_{NC}). Samples were considered positive if the S/P ratio was greater than 60%, doubtful when ranged between 50 and 60%, and negative when lower than 50%. Optical density (OD) was measured at 450 nm using the Cytation 3 microplate reader (BioTek).

Histology and histopathology

Nasal Turbinate, trachea and lung tissue samples stored in 10% buffered formalin were left overnight before embedding in paraffin blocks. Sections of 4 µm were cut, dewaxed and stained using hematoxylin and eosin. All tissues were scored independently by two veterinary pathologists. Lesions were graded by severity (mild, moderate, and severe) and distribution (focal, multifocal, or diffuse). Immunohistochemistry targeting hACE2 and SARS-CoV-2 nucleocapsid protein was performed on fixed tissue sections. Envision detection reagents were used for visualisation (Dako). Known hACE2 positive and negative sections were used to validate the assay.

QuPath implementation of pixel-wise H-score

QuPath (version 0.5.1) is an open-source software for digital pathology and whole slide image analysis [229]. The detailed explanation and calculation for the pixel-wise H-score are described elsewhere [230]. Briefly, the pixel-wise H-score is analogous to the traditional cell-based H-score but is applied to pixels as opposed to individual cells. Using the built-in pixel classifier in QuPath, a single user-defined threshold (1 for the

IHC images shown here) was set to detect all haematoxylin positive pixels and three separate thresholds, 0.2, 0.15 and 0.1 were set to detect and classify the DAB positive pixels (hACE2) as high, medium and low respectively.

Ethics statement

The generation and maintenance of transgenic pigs were approved by the Animal Welfare Ethical Review Body of the Roslin Institute and the infection study was approved by the Animal Welfare Ethical Review Body of the Moredun Research Institute. All experiments involving animals were authorized by the UK Home Office and were performed in accordance with the Animal (Scientific Procedures) Act 1986.

2.2.7 Statistical analysis

Microsoft Office Excel was the primary tool used for data analysis. All statistical analyses were performed using GraphPad Prism. Mean and standard deviation are shown in figures unless otherwise specified.

CHAPTER 3

Characterisation of porcine cell lines for coronavirus replication and transgene expression

3.1 Overview and Background

The majority of studies on IFN have focused on humans and model species, such as mice. While these studies have helped dissect the IFN response and contributed to our understanding of species-specific restriction factors, studies in other species, are vital to fully understanding the role of IFN as a barrier to virus spillover events. Virus infections of farmed animals are not only a major risk to food security but also underpin the risk of zoonotic events, due to their ability to act as intermediate hosts for zoonotic viruses and their close proximity to humans. To address this knowledge gap, we have generated arrayed ISG screening libraries in farmed animal species, including pigs, to complement and extend the studies already performed in human cells.

In this project, we are aiming to deliver porcine ISGs via a 2nd generation HIV-1-based lentiviral vector into a cell line of porcine origin which is permissive to the adapted PRCV and non-adapted SARS-CoV-2. As mentioned in the first chapter, PRCV has a limited impact on the pig industry and has therefore received much less research interest compared to other porcine viruses. Resources such as detection methods and handling are not well established and are mostly based on the closely related and well-studied TGEV. It has been reported that SARS-CoV-2 can replicate in some porcine cell lines but comprehensive studies regarding this are lacking. Despite the extensive use of HIV-1-based lentiviral vectors in various human cells, their use in cells of other animal origins is less common. Knowledge regarding their efficiency in porcine cell lines is also limited. In this chapter, we will characterise the replication of the two viruses in porcine cell lines and test the cell line efficiency for transduction.

Immortal cell lines are often used in biological research in addition to primary cells. Throughout the decades, researchers have developed a wide variety of cell lines which suit the needs and purposes of different research. The largest cell line depository, the American Type Culture Collection (ATCC) has over 4000 cell lines in its collection, with different species origins, tissue/disease types, and signalling pathways.

However, despite there being a large number of available cell lines, we sometimes still rely on suboptimal cell lines for virus research. The characteristics of viruses *in vivo* versus *in vitro* can vary greatly. Cell lines that are expected to be susceptible may not be, while cell lines that turn out to be susceptible may be from a clinically irrelevant species. Compromises are often needed when it comes to cell line selection. For example, some SARS-CoV-2 studies rely on the Vero cell line originated from the kidney of African green monkeys. Vero cells are susceptible to a wide range of viruses but their species origin means they are a less relevant model for human disease. Moreover, passaging SARS-CoV-2 in Vero cells can lead to adaptation in the virus genome, such as the loss of the S1/S2 furin-cleavage site, further undermining the clinical relevance of studies carried out in this cell line [231]. In contrast, cell lines such as Calu-3 (lung adenocarcinoma) and Caco-2 (colorectal adenocarcinoma) are of human origin. However, they generate lower virus titres and are highly heterogeneous, meaning they are a mixed population and results obtained may vary based on culturing methods [225]. Overexpression of entry receptors is another common approach to render non-susceptible cell lines permissive to virus infections. However, this is only applicable to viruses that have a well-defined functional receptor. Besides, there are limitations and drawbacks to the use of naturally non-permissive cell lines in virus research. Other commonly used cell lines such as hamster

kidney BHK-21 and human kidney HEK293, as well as Vero cells, have dampened IFN signalling pathways, making them less than ideal when the innate immune response is of interest in the study.

One can imagine this problem is even worse when it comes to viruses of farmed animals, as fewer resources, such as cell lines and knowledge on virus entry are available. PRRSV is another *Nidovirus* that severely impacts the pig industry, resulting in increased research interest. However, this virus has a highly restricted tropism in cells with monocyte or macrophage lineages. Apart from the use of primary macrophages from pigs which is inconvenient and highly heterogeneous, the most commonly used cell line for virus isolation and propagation remains the monkey-derived MARC-145, as there are very limited porcine macrophage cell lines [232]. There is also a lack of highly susceptible porcine cell lines for PEDV studies, and again Vero cells become the suboptimal solution as they are not the natural host and attenuation through passage was also observed [233, 234]. Overexpression of the entry receptor is unreliable in the case of PEDV as the actual receptor is yet to be identified despite years of effort [233]. Developing cell lines for PEDV and PRRSV study has therefore become an active field of research.

Apart from species origin, tissue type and signalling integrity, the efficiency of a cell line for transgene expression is also taken into consideration, especially for gene function studies. The two most common ways to artificially introduce foreign nucleic acid into a cell line are transfection and transduction. The former refers to the transfer of naked DNA or RNA by means other than viral infection, such as lipofection and electroporation; the latter refers to the utilisation of modified viral vectors such as

lentivirus and adenovirus. Some cell lines such as HEK293 are extremely popular in biological research due to their high transfection efficiency. In contrast, primary cells are notoriously difficult to transfect, and extensive optimisation and special reagents are often needed just to achieve a suboptimal efficiency. Transduction has therefore become a popular approach for gene expression in cells that are difficult to transfect, as the viral vectors can be modified to have a broader cell tropism. Nonetheless, disadvantages such as immunogenicity and cytotoxicity in transduced cells, as well as a lower packaging capacity compared to non-viral methods ($\sim$10 kb to $\sim$100kb) have limited the use of viral vectors in certain cells [235, 236].

The concept of viral-based methods for gene delivery was first described in 1983 with vectors based on a mouse retrovirus [237]. The most distinctive feature of viruses from the *Retrovirdae* family is that they can insert a DNA copy of their RNA genome into the DNA of the host cells, thereby permanently changing the genome of the cells and their progeny. This makes them ideal vectors for gene delivery as they can integrate a gene of interest (GOI) into the host genome for long-term stable expression. However, retroviral vectors have a disadvantage in that they transduce non-dividing or slowly dividing cells inefficiently, as they require the deconstruction of the nuclear membrane (which occurs during cell division) to gain access to the host genome for integration. In contrast, lentiviral (a genus under *Retrovirdae*) vectors can actively transport their core into the cell nucleus for integration and are therefore able to transduce cells irrespective of their division status [238]. All retroviruses require the basic genes *Gag* (encoding structural proteins), *Pol* (encoding reverse transcriptase, integrase, and other required enzymes) and *Env* (encoding the envelope glycoprotein) to form functional viral particles.

This knowledge has prompted the development of replication-incompetent vectors that can be used safely. HIV-1-based vectors are the most common lentiviral vectors and the 1st generation of vectors was made of three plasmids: **(1)** a plasmid encoding a heterologous envelope protein (e.g. G glycoprotein of VSV with broad cell tropism); **(2)** a packaging plasmid driven by a CMV promoter encoding all HIV genes (including non-essential accessory genes) except *Env* and *Vpu*; and **(3)** a transfer plasmid carrying a GOI and all other cis-acting element involved in virus integration, packaging and reverse transcription. Vectors produced by these plasmids do not contain any basic genes and therefore cannot further replicate in transduced cells [239]. However, the packaging plasmid retains a significant proportion of the HIV genome and has a potential risk of reverting to wild-type viruses. The 1st generation is therefore rarely used since the development of newer packaging systems, where accessory genes not essential for virus particle production were removed. Only *Tat* (enhancing viral gene transcription) and *Rev* (enabling nuclear export of viral transcripts) remained in the 2nd generation, with *Tat* removed and *Rev* provided on a separate packaging plasmid in the 3rd generation (therefore four plasmids instead of three) [240, 241].

The cross-species use of lentiviral vectors faces a similar challenge to the cross-species transmission of viruses. As discussed in the first chapter, HIV-1 has evolved and adapted to some human antiviral ISGs such as TRIM5α and tetherin, and therefore HIV-1-based lentiviral vectors can bypass these blocks and efficiently transduce human cells. However, the orthologs of these ISGs in other species can act as potent barriers to antagonize viruses and thus vectors [241]. The transduction efficiency of HIV-1-based vectors in cow-, rabbit- and mouse-derived cells are therefore significantly lower than

primate cells [242]. Although these restrictions can sometimes be saturated by using a higher titre, this will risk modifying the cell physiology and trigger unintended immune responses directed against the vectors [243, 244]. Vectors based on lentiviruses from other animals such as feline immunodeficiency virus, equine infectious anaemia virus and bovine immunodeficiency virus are also developed, although their use and availability are much less common compared to vectors based on primate lentiviruses [241, 245–247].

In this chapter, we have generated the required reagents such as cell lines and optimised protocols such as the transduction of pig cell lines. The growth kinetics of PRCV and SARS-CoV-2 have also been characterised. These work are essential for the screening experiments described in the next chapter.

3.2 Results

3.2.1 Detection and growth of PRCV in porcine cell lines

Previous screens using human- and avian-adapted IAV against a panel of human ISGs successfully identified a cross-species primate antiviral factor that specifically restricted avian IAV. Here, we use a similar approach to identify cross-species antiviral factors in pigs with the potential to restrict zoonotic coronaviruses. However, coronaviruses do not present a similar situation to IAV, with a range of highly related viruses that infect multiple species. Instead, an alternative screening approach was taken using a porcine-adapted respiratory coronavirus PRCV and human-adapted respiratory coronavirus SARS-CoV-2 in pig cells using the porcine ISG library.

PRCV had not been studied or used in the Grey laboratory before the start of this project and therefore no guidelines or protocols had been established. Materials including viruses (PRCV strain ISU-1 and AR310) and cell lines (ST; swine testis cells), were kindly provided by our collaborator at the Pirbright Institute. The master stock from Pirbright (P1) was propagated in ST cells using published methods for TGEV to generate P2 stocks. Briefly, cells were infected with PRCV diluted in serum-free media and supernatants were harvested 3 or 4 days post-infection, when complete cytopathic effect (CPE) was observed. However, the PRCV-infected ST cells show only minor CPE at 4 days post-infection. We initially hypothesize this is a phenotypic difference between PRCV and TGEV as it is a much less virulent virus *in vivo*. Regardless, the supernatant was harvested, aliquoted and stored at -80°until use.

Establishing effective detection methods to quantify virus replication would be key for the planned ISG screens. The S gene, ORF3 gene and N gene were selected as initial targets for PCR amplification, as these genes were used for the detection of the highly similar virus TGEV. At the time, ISU-1 was the only PRCV strain with a publicly available genome. Primers were therefore designed based on the reference genome and this strain has been selected for initial optimization. To test the specificity and efficiency of the designed primers, the P2 stocks were used to infect ST cells seeded in 24 well plates. As we have not established a protocol for titring PRCV at this stage, cells were infected at an arbitrarily higher multiplicity of infection (MOI; 1/5 of P2 stock). 24 hours post-infection (HPI), the cell lysate from infected ST cells was harvested using TRIzol and RNA and cDNA samples were prepared. A 10-fold dilution series (10^{-1}-10^{-5}) of the samples was prepared and added to a dye-based reverse transcription-quantitative polymerase chain reaction (RT-qPCR) using 500nM of primers targeting the S, ORF3 or N gene. The raw cycle threshold (CT) value was recorded and the efficiency of each primer pair was calculated as follows:

$$\text{Efficiency (\%)} = (10^{\frac{-1}{\text{Slope}}} - 1)\text{x}100$$

The perfect primer pair would have an efficiency of 100%, doubling the number of molecules of the target sequence during each reaction cycle. However, this is not normally achievable in practice and pairs with an efficiency between 90% and 110% are considered ideal. Primers with low efficiency usually suggest they do not bind to the target sequence efficiently, resulting in less than twice the molecules produced every cycle. In contrast,

primers with over 100% efficiencies could suggest they are binding to more than 1 target, resulting in unspecific amplification and more than twice the molecules produced each cycle.

The primer pair targeting the S gene had an efficiency of nearly 200% and gave the highest CT, suggesting it might be producing unspecific products and the designed primers are sub-optimal. In contrast, the ORF3 and N primer pairs produced similar results with efficiencies close to 100%, suggesting the designed primers would work well (**Fig 3.1**). As the efficiency of the N primer pair is slightly closer to 100%, the N gene has been selected for PRCV detection by RT-qPCR hereafter.

As dye-based RT-qPCR has a higher risk of producing inaccurate results as binding to any dsDNA including unspecific products and primer dimers can occur, a TaqMan probe targeting a region of the amplicon of the above PRCV N primer pair was also designed. To test the efficiency of the probe-based assay, a plasmid control encoding the expected amplicon was serially diluted and a known copy number of plasmids was used as input for RT-qPCR. As expected, the assay showed sufficient efficiency when paired with the plasmid control and was used to generate a standard curve for absolute quantification of PRCV N copy number in samples hereafter (**Fig 3.2A**).

The majority of published studies involving PRCV or TGEV used ST cells as they are highly susceptible. This cell line has therefore been selected as the control and reference cell line for PRCV infection in this project. In addition to ST cells, we have access to other porcine cell lines including newborn pig trachea (NPTr), newborn swine kidney (NSK) and porcine kidney 15 (PK15). These cell lines, along with four other cell lines

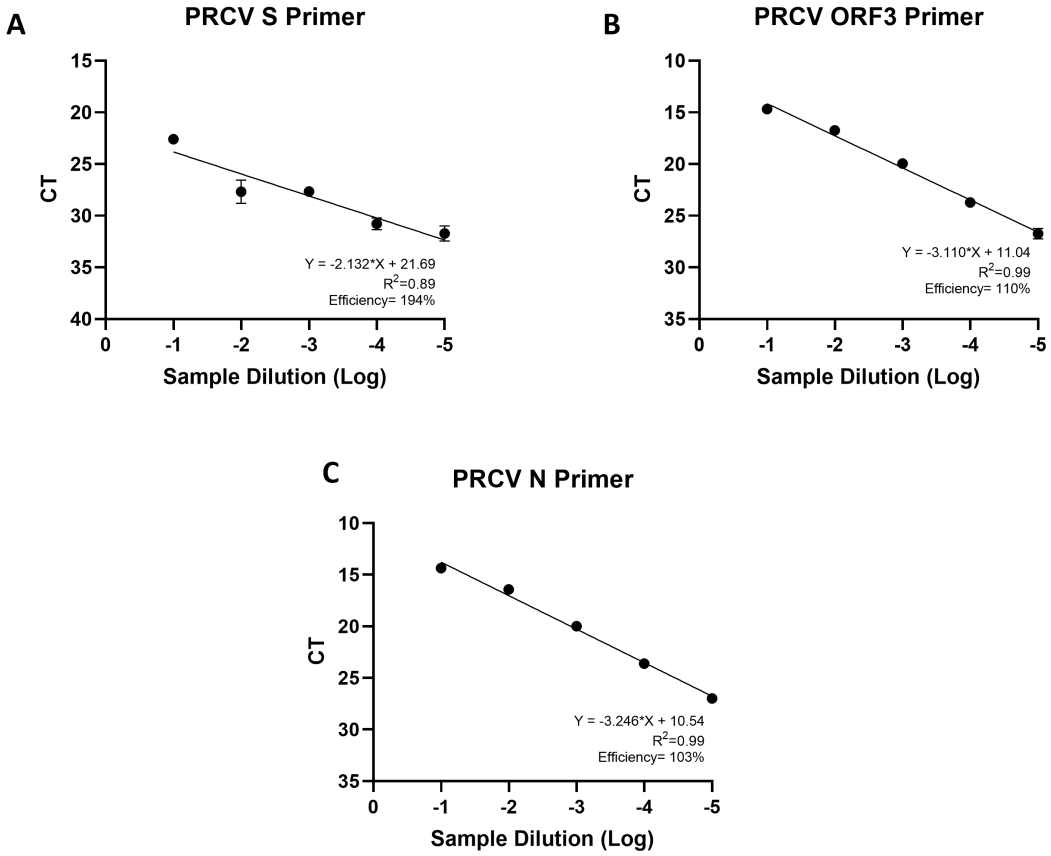


Figure 3.1: **Primer efficiency for PRCV detection**. Three primer pairs were designed based on the published sequence of PRCV strain ISU-1. The cDNA sample from cells infected at an arbitrarily higher MOI (1/5 of P2 stock with unknown titre) was subjected to 10-fold serial dilution and used as templates for dye-based RT-qPCR (SYBR green). The raw CT and efficiency of the primer pairs targeting **(A)** the spike (S) gene, **(B)** open-reading frame 3a (ORF3a) and **(C)** nucleocapsid (N) gene were shown. Linear regressions were calculated to assess reaction efficiency from the slope. Results represent the average of three technical replicates. Error bars represent the standard deviation. CT= cycle threshold

of human origin (MRC-5, Caki-1, Calu-1 and BEAS-2B), were seeded in 24-well plates and infected at an arbitrarily higher MOI with PRCV ISU-1 to test relative susceptibility. At 24 HPI, the cell lysate and supernatant of infected cells were harvested and the N copy number in the samples was determined.

ST and NPTr cells are both highly susceptible and permissive, as a large amount of N gene is detected both within the cell and in the supernatant. This suggests PRCV is actively replicating and producing virus particles in these cells. The NSKs are also susceptible but to a lesser extent, as the N copy number in the cell lysate and supernatant is approximately 1 log and 2 log lower than ST and NPTr respectively. No copies of N were detected in infected PK15 cells. The N copy number in cell lysate samples is expected to be higher than that in the supernatant, as the subgenomic RNA produced during PRCV replication in the cells contains the genomic sequence of N (**Fig 3.2B and C**). As expected, none of the human cell lines were susceptible to PRCV and no N gene was detected. The spike protein of TGEV cannot recognise the human APN receptors and it is very likely that the truncated spike protein of PRCV cannot recognise human APN either [248].

We next sought to characterize the growth kinetics of PRCV ISU-1. As NPTr and ST cells were shown to have the highest susceptibility, they were selected for further experiments and were infected at an arbitrarily lower MOI (1/50 of P2 stock with unknown titre). RNA viruses such as IAV can produce progeny virus within hours, while some DNA viruses such as HCMV can take days or even weeks. PRCV is therefore expected to have a rapid life cycle and the time course experiment was designed accordingly. Cell lysate

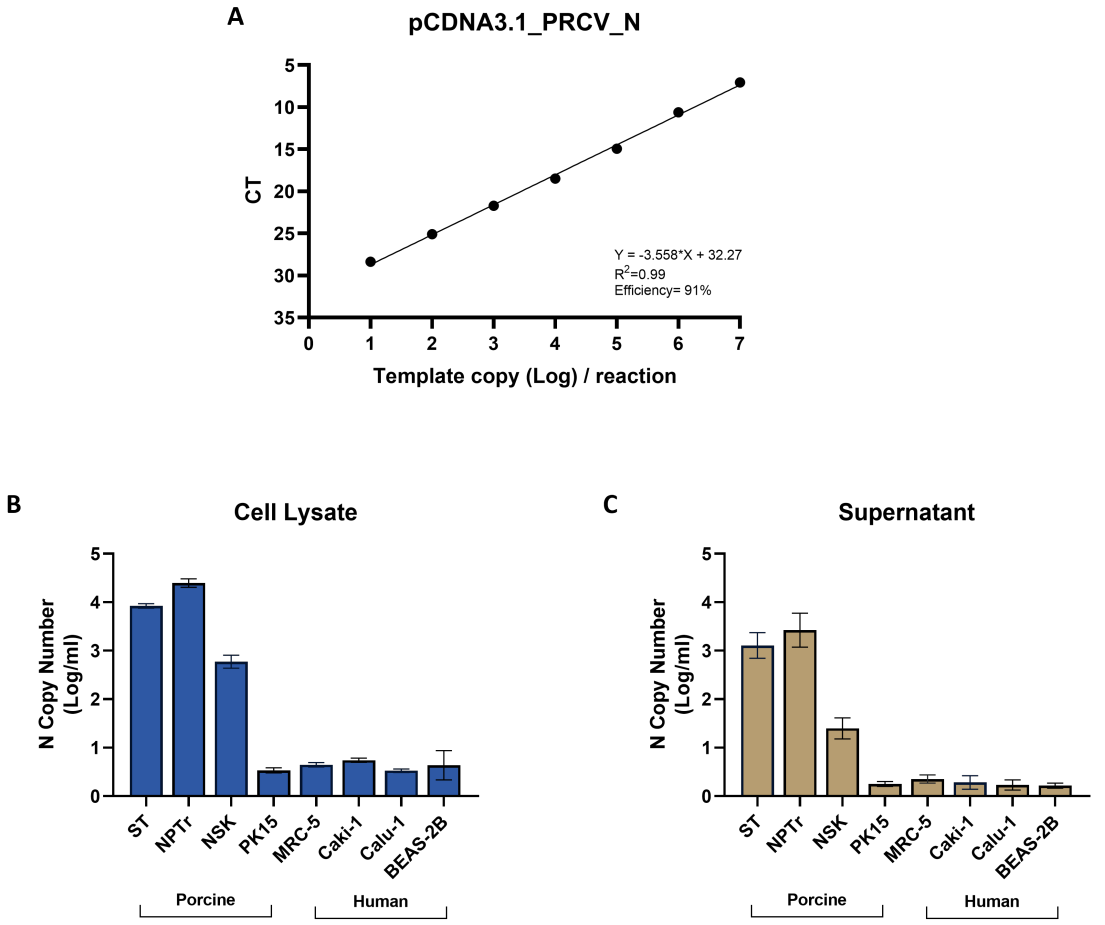


Figure 3.2: **PRCV ISU-1 N copy number in various cell lines**. **(A)** A standard curve was generated using a plasmid control encoding the expected amplicon of a designed TaqMan assay targeting the N gene of PRCV. Selected porcine and human cell lines were infected at an arbitrarily higher MOI (1/5 of P2 stock with unknown titre) and harvested at 24 HPI. The N copy number of **(B)** cell lysate samples and **(C)** supernatant samples was calculated using the standard curve. A new standard curve was generated for each RT-qPCR run. Linear regressions were calculated to assess reaction efficiency from the slope. Results represent the average of three independent experiments, each with three technical replicates. Error bars represent the standard deviation. CT= cycle threshold

and supernatant samples were harvested at 2, 8, 12, 16, 24, 34 and 48 HPI and the N copy number were determined using the standard curve generated by the plasmid control. Even at a lower MOI, the N copy number in the cell lysate peaked at approximately 8 HPI for both cell lines. After 8 hours there was little increase in the N copy number. The slight decrease of N copy number in ST cells from 24 HPI onwards might be due to the minor CPE observed at 24 HPI, which may lead to the release of the subgenomic RNA. No CPE was observed in PRCV-infected NPTr. The N copy number in supernatant, in contrast, gradually increased in the first 24 hours and then plateaued, suggesting virus production peaked at 24 HPI (**Fig 3.3**). We therefore picked 24 hours as the time of harvest for the following experiments.

We have established an RT-qPCR-based assay for the detection of PRCV. However, the measurement of the N copy number does not reflect infectious virus titre and the MOI we have been applying in previous experiments is only relative. A different detection method such as antibodies can be used to measure viral protein production and virus spread. There are multiple commercially available antibodies targeting TGEV and other related coronaviruses but to the best of our knowledge, none claim to have reactivity towards PRCV. We initially expected antibodies targeting TGEV would cross-react with PRCV as this has been reported to occur *in vivo* (see **1.1.3**). We have tested 3 antibodies and 1 antiserum that are reported as reactive against TGEV for western blot, flow cytometry and fluorescence microscopy. However, signals were only detected by fluorescence microscopy and full staining of cells was not achieved. An increased amount of signal was observed in samples infected with an arbitrarily higher MOI (1/5 of P2 stock with unknown titre) at 24HPI (**Fig 3.4**). Nonetheless, signals were only

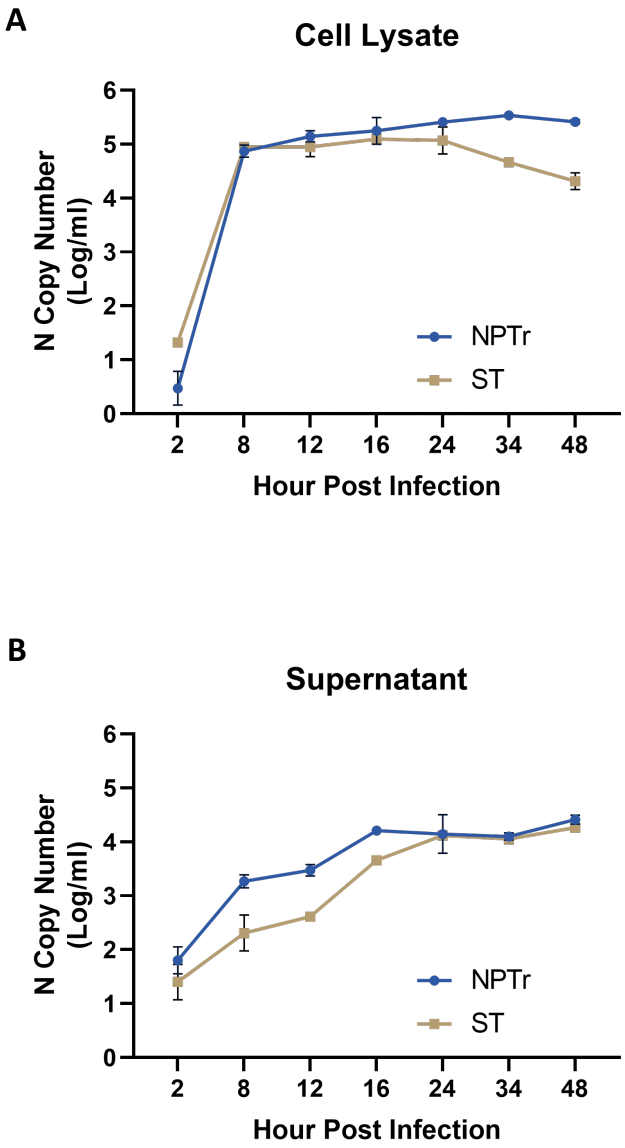


Figure 3.3: **Growth Kinetics of PRCV in ST and NPTr cells**. ST and NPTr cells were infected by PRCV ISU-1 at an arbitrarily lower MOI (1/50 of P2 stock with unknown titre). Samples were harvested at designated time points. The N copy number in **(A)** cell lysate and **(B)** supernatant samples was determined by a standard curve generated by a control plasmid. Results represent the average of two independent experiments, each with two technical replicates. Error bars represent the standard deviation.

detected in a relatively small number of cells despite the initially thought high MOI infection. We speculate despite a significant increase in the N copy number detected from infected ST and NPTr cells supernatant samples in the first 24 hours (**Fig 3.3B**), only a small amount of infectious virus particles were produced. The titre of the P2 ISU-1 stock might be substantially lower than originally calculated. A method detecting infectious virus particles is therefore needed for the determination of actual virus titre.

The 50% tissue culture infectious dose (TCID50) on ST cells is a common approach for titring TGEV. We have therefore applied this method to determine the titre of the P2 ISU-1 stock. Consistent with observations during stock production, CPE associated with infection of ST cells with PRCV was minor and was not comparable to the complete CPE observed in TGEV-infected cells as reported in other studies. In conjunction with RT-qPCR assay and antiserum detection, we concluded the produced stock had a very low titre, with less than 10^3 TCID50/ml. A new P2 stock was produced using strain AR310 and a similar titre was obtained. We suspected both original P1 stocks we received had a low titre. Indeed, titring using TCID50 assays confirmed the stocks of both PRCV strains had only a slightly higher titre of 10^4 TCID50/ml.

The consistently low titres suggested a potential problem with the initial stocks of virus. As an alternative approach, PRCV strain AR310 was acquired directly from ATCC (P0). In contrast to the previous stocks of virus, infection of ST cells with the virus acquired from ATCC resulted in complete CPE, confirming the hypothesis that the original stocks of virus were attenuated (**Fig 3.5C and D**). All experiments described hereafter were performed using a P3 working stock generated by infection of ST cells from the ATCC-

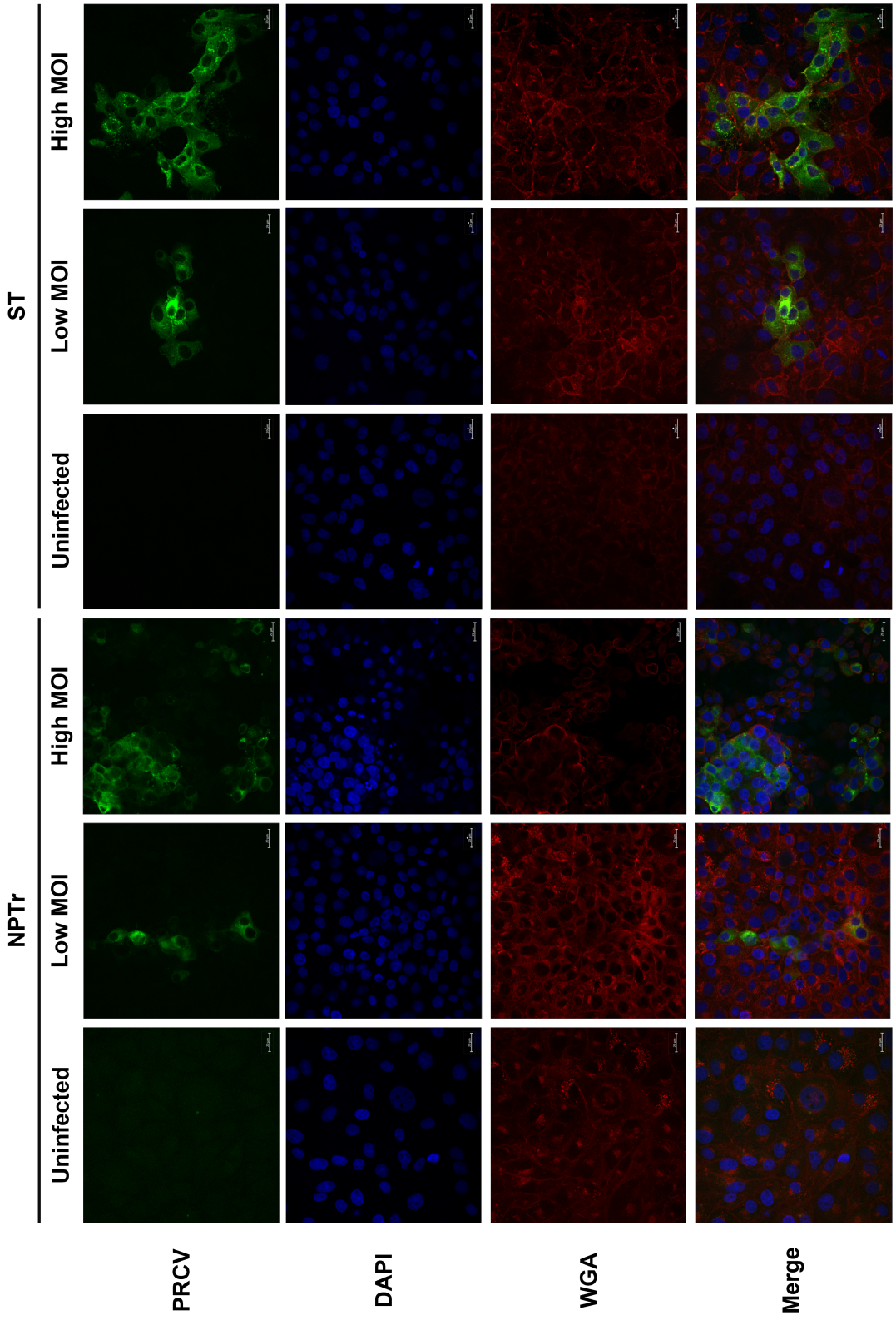


Figure 3.4: **Detection of PRCV ISU-1 by fluorescence microscopy**. ST and NPTr cells were infected at a relatively high (1/5) or low (1/50 of P2 stock with unknown titre) MOI and were fixed and permeabilized at 24 HPI. A polyclonal antiserum against TGEV was used for PRCV detection (Green), wheat germ agglutinin (WGA) and DAPI for membrane (Red) and nucleus (Blue) counterstain. Scale bar= 20µm. Image representative of 2 independent experiments.

acquired PRCV AR310 virus.

To measure the production kinetics of infectious virus particles from different porcine cell lines, we have performed a similar experiment as shown in **Fig 3.3**, except the supernatant samples were subjected to TCID50 on ST cells instead of measuring viral RNA levels. Consistent with previous experiments, where N copy numbers were measured, ST and NPTr cells remain the most susceptible cell lines, producing more than 10^6 TCID50/ml at 24 HPI. The titre increased most significantly between 8 and 24 HPI but slowed after 24 HPI. The amount of infectious virus particles produced from NSK is higher than previously observed, based on the N copy number in NSK supernatant in previous experiments (**Fig 3.2C**). The titre continued to increase throughout the remaining time course, reaching levels matching ST and NPTr cells. In contrast, PK15 cells did not produce any virus throughout the entire infection course (**Fig 3.5A**).

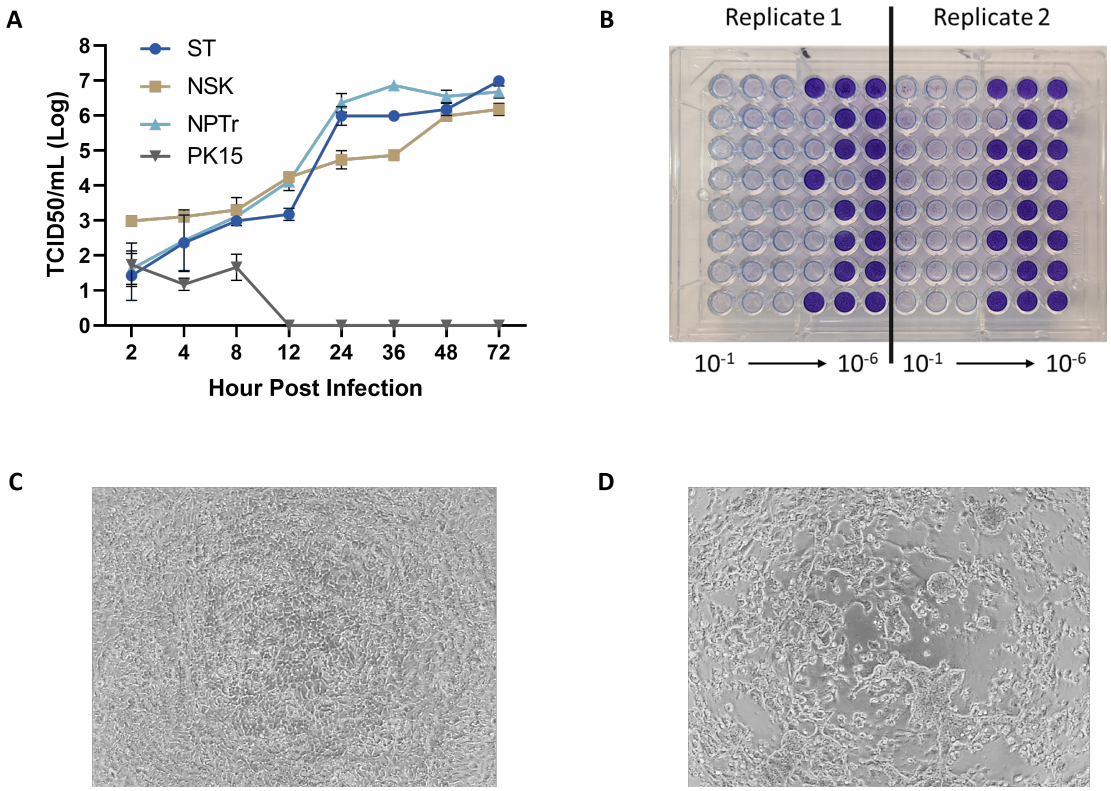


Figure 3.5: **Measuring PRCV infectious titre by TCID50**. **(A)** Selected porcine cell lines were infected at a MOI of 0.01. Supernatant samples were harvested at indicated time points and subjected to TCID50 on ST cells. **(B)** Example image of TCID50. Samples were 10-fold diluted in duplicate and used to infect ST cells in 96-well plates. Transparent wells represent cell death resulting from PRCV infection and stained wells represent surviving wells. Titres were calculated using the Spearman–Kärber method. Representative images were taken at 72 HPI for **(C)** uninfected and **(D)** PRCV-infected ST cells. Results represent the average of two independent experiments, each with two technical replicates. Error bars represent the standard deviation.

3.2.2 Generation and characterisation of cell lines overexpressing porcine APN

To conduct the lentiviral ISG screen against PRCV, the selected porcine cell line needed to be highly susceptible to PRCV but also transducible with the lentiviral vector used for the screens. Previously, we cloned the open reading frames (ORFs) of 432 porcine ISGs (and green fluorescence protein (GFP) as a negative control) into the destination vector pSCRPSY (GenBank accession KT368137) via gateway recombination cloning. Briefly, this vector consists of a minimal packageable HIV-1 genome containing all essential cis-acting sequences. The HIV-1 5' long terminal repeat (LTR) has been partially deleted and fused to a CMV promoter for increased safety. The *Gag*, *Pol*, *Vif* and *Vpr* genes were deleted and replaced with a Puro-2A-TagRFP cassette for the expression of puromycin N-acetyltransferase (PAC; for puromycin resistance) and Tag red fluorescence protein (TagRFP) reporter via a foot-and-mouth-disease virus 2A stop-start peptide. The *Vpu* and *Env* genes are also inactivated or deleted (**Fig3.6A**). Cells transduced with pSCRPSY vector express HIV-1 *Tat* and *Rev* along with the inserted cDNA from completely spliced 'early' HIV transcripts. The Puro-2A-TagRFP cassette is only expressed from unspliced 'late' transcripts, as unspliced transcripts cannot be exported from the nucleus for translation without the expression of *Rev*. The design of this plasmid aims to minimize the chance of a cell expressing the transduction marker (puromycin resistance and TagRFP) without expressing the inserted cDNA (ISG or GFP) [249].

Our group has experience using the pSCRPSY vector to transduce primary human

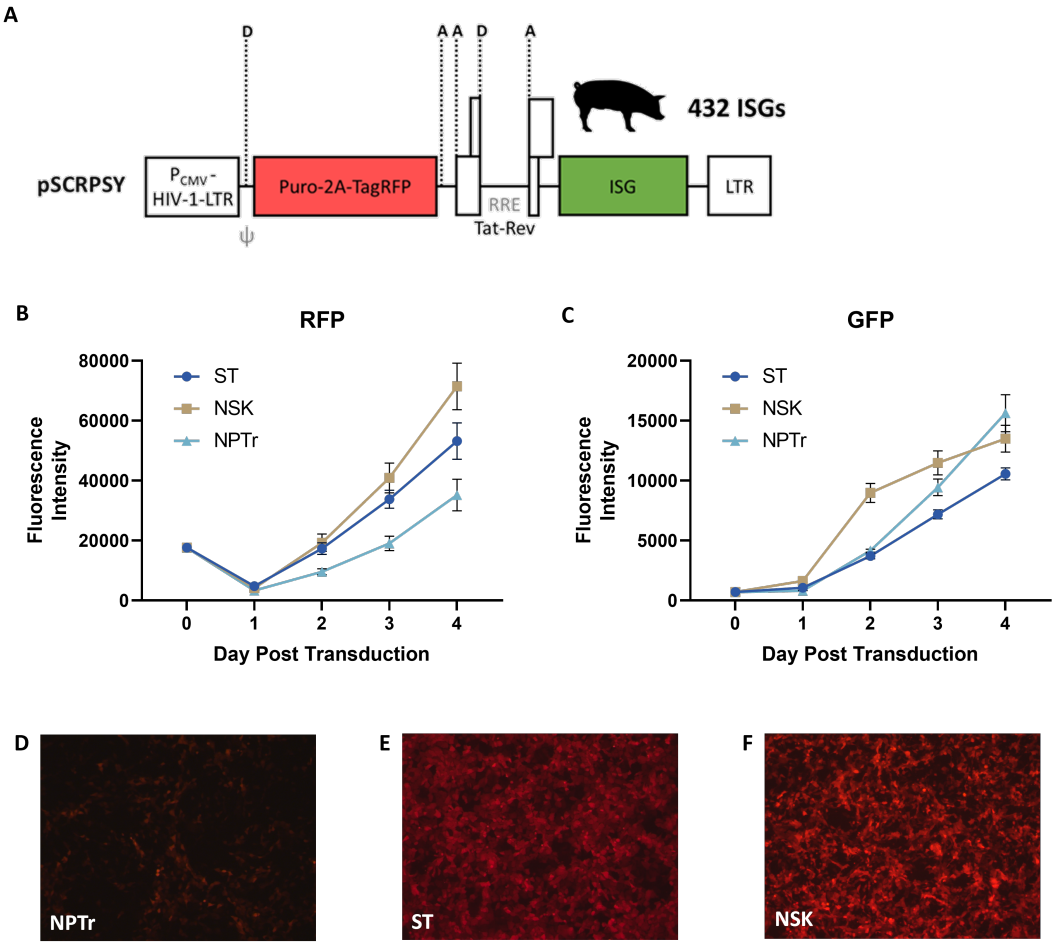


Figure 3.6: **Transduction efficiency of PRCV-susceptible cell lines.** **(A)** Schematic representation of the pSCRPSY plasmid. Indicated cell lines were transduced by lentivirus generated from pSCRPSY-GFP and fluorescence intensity of **(B)** RFP and **(C)** GFP were recorded daily for 4 days. Representative image of RFP expression in transduced **(D)** NPTr, **(E)** ST and **(F)** NSK cells at 3 days post-transduction. Results represent the average of three independent experiments, each with three technical replicates. Error bars represent the standard deviation.

fibroblast (NHDF) cells for screening experiments [250]. NPTr cells were selected for pilot studies as they were the most susceptible cell lines for PRCV infection in previous experiments and are of respiratory tract origin, the natural replication site of PRCV. However, it was soon found that the transduction efficiency (using TagRFP fluorescence intensity as a surrogate) of NPTr cells is substantially lower than that of NHDF cells. The amount of lentivirus used to transduce NPTr cells was increased in an attempt to enhance transduction efficiency but was unsuccessful. We therefore tested the transduction efficiency of ST and NSK cells.

ST, NSK, NPTr and NHDF cells were transduced with the GFP pSCRPSY control lentivirus and the fluorescence intensity of RFP and GFP was measured by plate fluorometry. Of the three susceptible cell lines, NPTr cells demonstrated the lowest RFP signal, only around 50% of ST and one-third of NSK cells at 4 days post-transduction(**Fig3.6B**). However, the GFP signal from NPTr is similar to that of NSK at 3 and 4 days post-transduction, with ST cells giving the weakest signal (**Fig3.6C**). Images taken at 3 days post-transduction by fluorescence microscope are consistent with the RFP signal measurement, with few RFP-positive cells present in the transduced NPTr cells, in contrast to ST and NSK cells (**Fig3.6D, E and F**). The TagRFP in pSCRPSY is a marker for transduction whereas expression of GFP directly correlates with ISG expression, and could therefore be argued to be a more important indicator of the likely success of a screen. The above data shows that, despite the porcine cell lines having differential RFP signals, the GFP signals are similar and their ability to express the transgene should, therefore be similar. Nonetheless, further experiments suggest NSK is the only suitable porcine cell line we have access to for the screening experiment (further discussed in Chapter

4). It is currently unclear why the different porcine cell lines produce relatively similar GFP signals but divergent RFP signals when transduced with the pSCRPSY lentiviral vector. One possible explanation is the varying efficiency of HIV-1 accessory proteins in the porcine cell lines, as RFP expression depends on the function of *Rev* while GFP expression is independent.

The susceptibility of a cell line towards a virus can be altered by overexpressing the relevant viral entry receptor, but it is much more difficult to adjust its transduction efficiency. It is, therefore, more reasonable to increase the susceptibility of NSK cells towards PRCV than to attempt to increase the transduction efficiency of NPTr cells. The higher transduction efficiency exhibited by NSK cells also allows easier transgene expression by lentiviruses. The entry receptor of PRCV has been identified as porcine APN (pAPN). We hypothesized that over-expression of pAPN in NSK cells would enhance susceptibility to PRCV. In addition to generating a porcine cell line more susceptible to PRCV, one of the major aims of this project is to compare the effect of porcine and human ISGs against coronaviruses. A human cell line susceptible to PRCV infection would, therefore, be valuable for comparative screening, potentially allowing the identification of species-specific ISG restriction factors. As NHDF cells had previously been successfully used for screening experiments, we overexpressed pAPN to test if it was sufficient to render NHDF cells susceptible to PRCV. Briefly, pAPN cDNA from NPTr cells was cloned into a pLenti-CMV-Blast plasmid vector (Addgene #17486). An HA tag was added to the carboxy terminus of pAPN, as commercially available antibodies cannot differentiate between pig and human APN. Both wild-type NSK and NHDF cells were transduced with the pAPN lentivirus to generate pAPN-NSK and pAPN-NHDF

cells.

The resulting replication analysis showed that overexpression of pAPN enhanced the susceptibility of NSK cells to PRCV infection. The N copy numbers in the cell lysate and supernatant of pAPN-NSK were significantly higher than the parental cell line (**Fig3.7A and B**). TCID50 results also demonstrate a significant increase in infectious virus titre from 24 HPI onwards, reaching 10^6 TCID50/ml, similar to ST and NPTr cells (**Fig3.7E**). The expression of pAPN in wild-type NSK cells is lower than that in ST cells at the mRNA level, potentially explaining the higher susceptibility in the latter (**Fig3.7C**). In contrast, overexpression of pAPN in NHDF cells did not increase susceptibility, as there is no difference in N copy number from the cell lysate and supernatant samples between wild-type and pAPN-NHDF cells. The high level of pAPN mRNA in the stable cell lines suggests this is not a defect from inefficient transduction.

Previous studies have shown that exogenous expression of pAPN in Vero, BHK-21 and HeLa (human epithelial cell line) cells results in susceptibility to TGEV infection [251, 252]. Interestingly, Vero and BHK-21 cells are known to have defects in their IFN signalling pathways as stated above and some reports suggest HeLa cells do not respond well to IFN. In contrast, the NHDF cells we used are primary fibroblast cells with an intact IFN system. We hypothesised that the intact IFN system in NHDF cells contributed to the restriction of PRCV in pAPN-NHDF cells despite the overexpression of entry receptors. To test this hypothesis, we first aimed to reproduce the results obtained from the above studies by generating Vero cells stably expressing pAPN. The protein expression of pAPN was confirmed by western blot for the HA tag in pAPN-Vero cells (**Fig3.7B**).

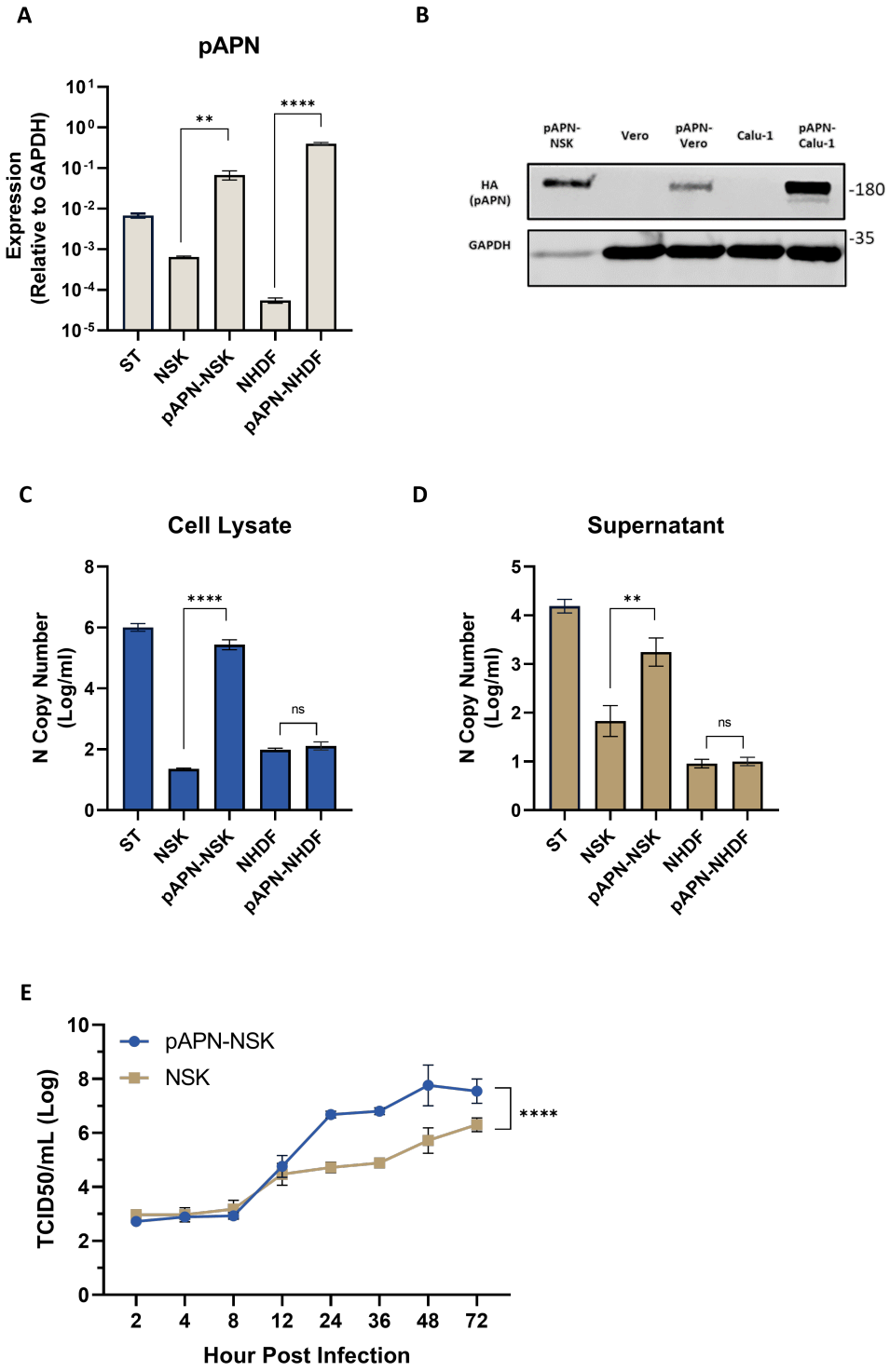


Figure 3.7: **PRCV susceptibility of cell lines overexpressing porcine APN.**

Figure 3.7: **PRCV susceptibility of cell lines overexpressing porcine APN.** NSK and NHDF cells stably expressing porcine APN (pAPN) were generated by lentiviral transduction. **(A)** The pAPN levels of the stable cell lines were measured by RT-qPCR. Expression levels were relative to GAPDH. **(B)** Protein expression of pAPN (through HA tag) was confirmed in additional generated cell lines. Indicated cells were infected by PRCV at MOI 0.1. **(C)** Cell lysate and **(D)** supernatant were harvested at 24 HPI and N copy numbers in samples were determined against a standard curve generated by plasmid control. **(E)** Growth kinetics of PRCV in NSK and pAPN-NSK. Cells were infected by PRCV at MOI 0.1, supernatants were harvested at indicated time points and measured by TCID50. Results represent the average of three independent experiments, each with three technical replicates. Error bars represent the standard deviation and unpaired t-tests were used to calculate statistical significance for **(A)**, **(B)** and **(C)**. Results represent the average of two independent experiments, each with two technical replicates. Error bars represent the standard deviation and two-way ANOVA was used to calculate statistical significance for **(E)**. ****= $p<0.0001$, **= $p<0.01$, ns= not significant.

Despite this, they were not susceptible to PRCV as no increase in N copy number was observed in the cell lysate or supernatant samples of infected cells at 48 HPI (data not shown). An aliquot of TGEV strain Miller (TC-adapted; purchased from ATCC) was acquired from our colleague and used to challenge the pAPN-Vero but we still failed to reproduce the published results. We repeated the experiment in the highly transducible Calu-1 human cell lines. However, no TGEV or PRCV replication was detected despite a higher pAPN protein expression level, suggesting the lack of susceptibility observed is not due to insufficient expression level. To better determine if the human IFN response indeed plays a role in the lack of susceptibility in pAPN-NHDF cells, the experiment can be repeated in IRF3 or STAT1 KO NHDF cells (both readily available). This will be more relevant as the cell lines used here are either not of human origin (i.e. Vero) or a human cell line with no known IFN-related deficiency (i.e. Calu-1).

Although we failed to generate a human cell line susceptible to PRCV for comparative screening, the pAPN-NSK cells did show increased susceptibility and could therefore be used for proposed ISG screening experiments. We next sought to test if these cells

have a competent IFN response and if PRCV is sensitive to Type I IFN treatment in general. Some ISGs included in our library are components of IFN and related signalling pathways and screening them in IFN-incompetent cells may overlook their effect. ISG20 was selected as the target gene for measuring IFN response as it has been reported to be robustly induced following IFN treatment or virus infection in a wide range of cell types and species. We initially treated the porcine cell lines with human IFNβ but no induction of ISG20 was observed. However, a recombinant human IFNα exhibiting bioactivity across multiple species (universal type I IFN) induced ISG20 expression in ST and pAPN-NSK cells by 24- and 6-fold respectively (**Fig3.8A**). This suggested both cell lines have an intact signalling pathway and are responsive to IFN treatment. IFN-treated cells were then challenged with PRCV at an MOI of 0.01 to test if viral replication was restricted. Results show that PRCV levels in the supernatant were reduced by more than 300- and 30-fold in ST and pAPN-NSK cells respectively (**Fig3.8B**). This difference is consistent with the observed difference in ISG20 induction, where ST cells potentially elicit a stronger IFN response and restrict PRCV replication to a larger extent.

Universal Type I IFN is a recombinant protein, which does not exist in nature and may trigger a different IFN profile in pig cells. Indeed, Shaw *et al.* showed that the up- and down-regulated genes in porcine primary fibroblast by universal IFN and porcine IFNα are similar but not identical [154]. PRCV may be able to overcome the response triggered by the natural porcine IFN but not by the artificial recombinant IFN. To compare the effect of the two IFNs on PRCV replication, ST and pAPN-NSK cells were treated with either universal or porcine IFNα at various concentrations. The supernatant RNA was extracted at 24 HPI and PRCV levels were compared to untreated samples. The half-

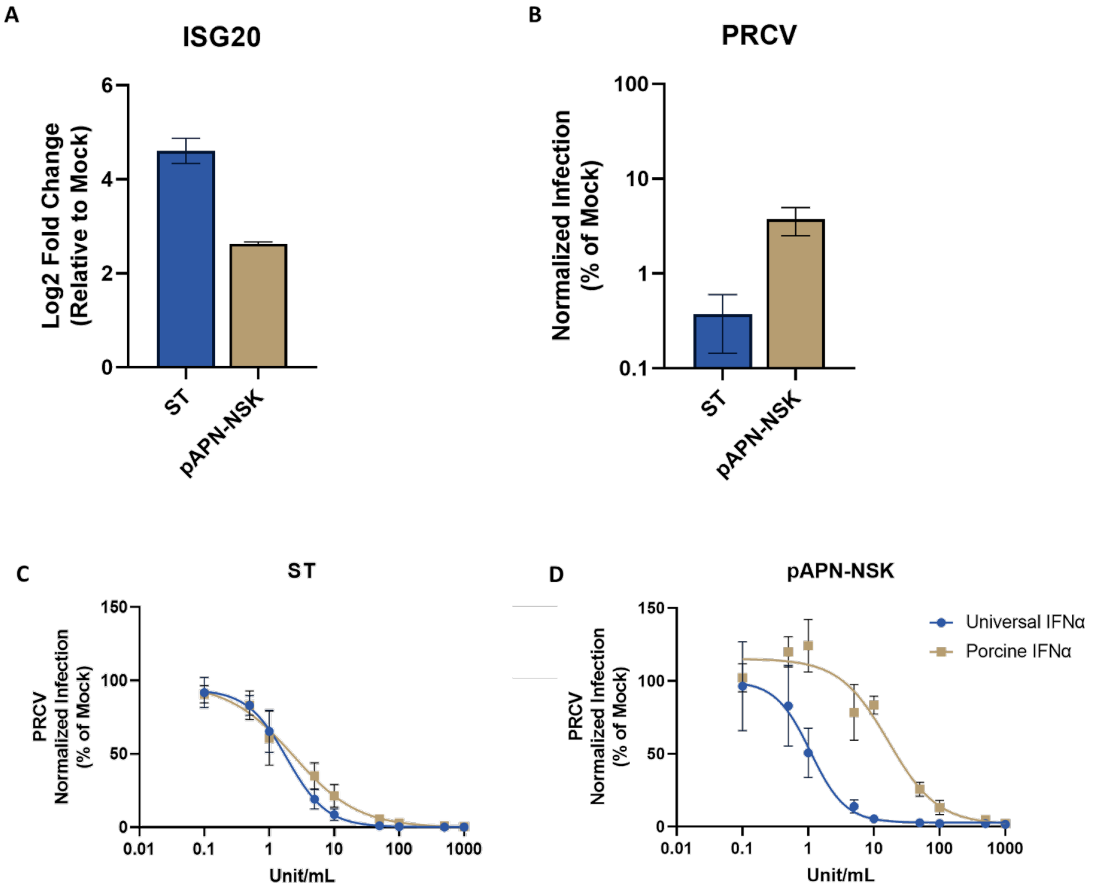


Figure 3.8: **Effect of Type I IFN on PRCV replication in ST and pAPN-NSK cells**. **(A)** ST and pAPN-NSK cells were mock-treated or treated with 1000U/ml universal type I IFN. Cell lysate samples were harvested at 24 hours and the expression of porcine ISG20 was determined by RT-qPCR. **(B)** ST and pAPN-NSK cells were mock-treated or treated with 1000U/ml universal type I IFN for 24 hours before being infected by PRCV at an MOI of 0.01. Supernatant samples were harvested at 24 HPI and the relative infection level was determined by RT-qPCR and the 2^(-delta CT) method. **(C)** ST and **(D)** pAPN-NSK cells were mock-treated or treated with the indicated amount of universal type I IFN or porcine IFNα for 24 hours before being infected by PRCV at an MOI of 0.01. Supernatant samples were harvested at 24 HPI and the relative infection level was determined by RT-qPCR and the 2^(-delta CT) method. Results represent the average of two independent experiments, each with three technical replicates and error bars represent the standard deviation for **(A)** and **(B)**. Results represent the average of three independent experiments, each with three technical replicates and error bars represent the standard error of the mean for **(C)** and **(D)**.

maximal inhibitory concentration (IC50) for both IFNs in ST cells is similar, ranging from around 1 to 5 unit/ml. Interestingly, the IC50 in pAPN-NSK cells is substantially different, while the universal IFN results in an IC50 of 1 to 3 unit/ml, the IC50 of porcine IFN increased by nearly 10-fold, ranging from around 10 to 30U/ml (**Fig3.8C and D**). This suggests the IFN responses triggered by both IFNs are more similar in ST than in pAPN-NSK cells, and PRCV seems to be better at countering the response elicited by porcine IFNα than universal IFN in pAPN-NSK cells. Nonetheless, the generated pAPN-NSK cells are generally IFN competent, highly susceptible to PRCV and transducible, making them ideal for ISG screening experiments.

3.2.3 Characterising SARS-CoV-2 infection of hACE2-NSK cells

While SARS-CoV-2 was reported to infect a wide range of animals, they cannot infect pigs *in vivo* despite *in vitro* evidence showing that the viral spike proteins can bind pig ACE2 (pACE2) [253]. Meanwhile, some studies have reported certain porcine cell lines such as PK15 and ST cells can support SARS-CoV-2 replication *in vitro* [222]. The discrepancy in susceptibility led us to hypothesise that apart from receptor compatibility, the IFN response may also play a role. We have therefore decided to apply the porcine ISG library to SARS-CoV-2 and compare it to PRCV to test the hypothesis. As we have overexpressed the entry receptor pAPN in NSK for PRCV, for consistent comparison, we have also overexpressed human ACE2 (hACE2) for SARS-CoV-2 by lentiviral transduction, generating hACE2-NSK cells. The susceptibility of these cell lines towards SARS-CoV-2 was then measured.

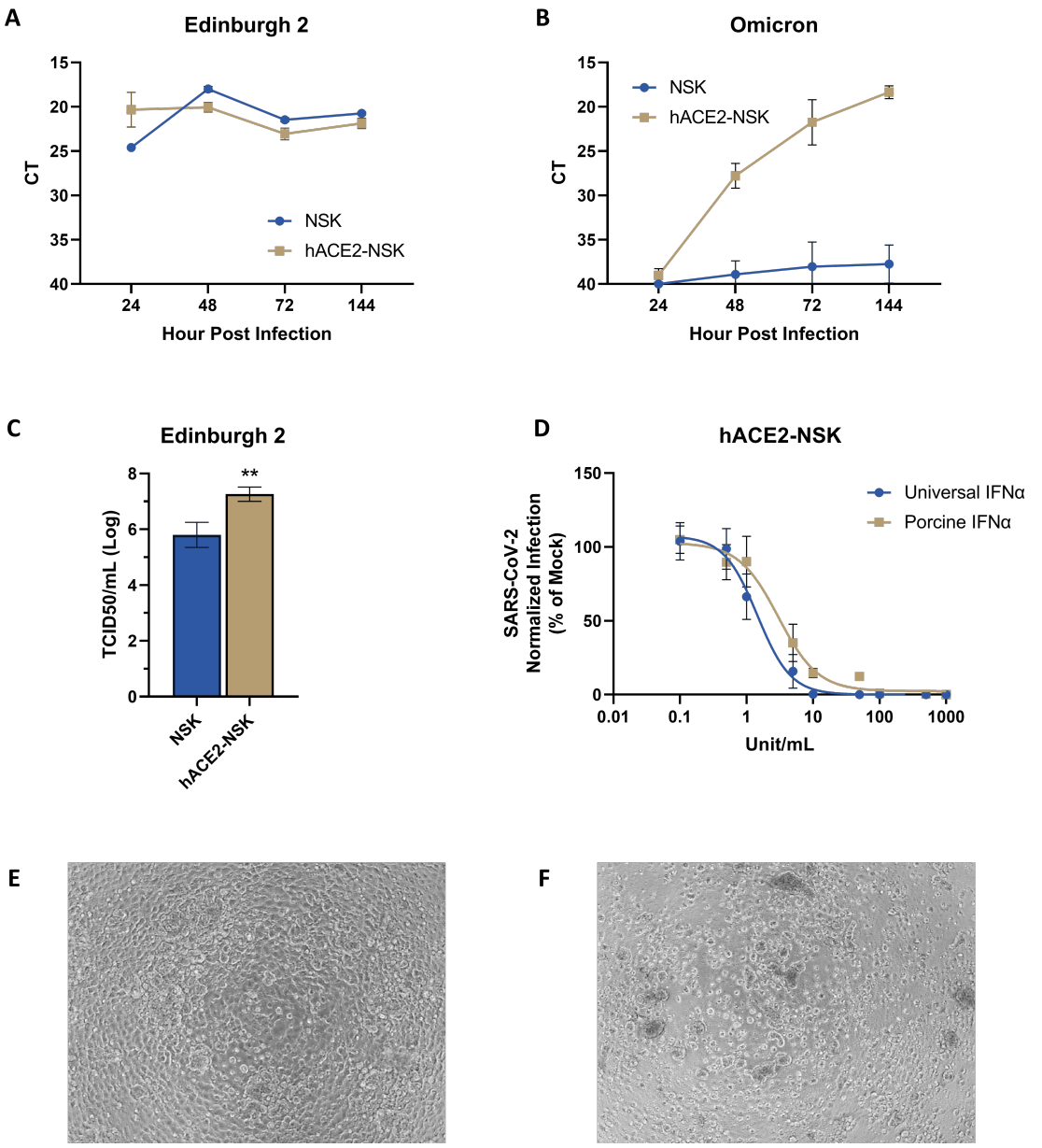


Figure 3.9: **SARS-CoV-2 replication and phenotype in hACE2-NSK cells.** Wild-type and human ACE2 (hACE2) overexpressing NSK cells were infected by SARS-CoV-2 **(A)** isolate Edinburgh 2 (EDB2) or **(B)** the Omicron variant at an MOI of 0.01. Supernatants were harvested at the indicated time points and RNA was extracted for dye-based RT-qPCR using the CDC N3 primer set. The raw CT values were reported. **(C)** Titre of EDB2 at 24 HPI measured by TCID50 on hACE2-NSK cells. **(D)** hACE2-NSK cells were mock-treated or treated with the indicated amount of universal type I IFN or porcine IFNα for 24 hours before being infected by EDB2 at an MOI of 0.01. Supernatant samples were harvested at 24 HPI and the relative infection level was determined by RT-qPCR and the 2^(-delta CT) method. Representative image of **(E)** uninfected and **(F)** EDB2-infected hACE2-NSK cells at 48HPI. Results represent the average of two independent experiments, each with three technical replicates and error bars represent the standard deviation for **(A)** and **(B)**. Results represent the average of three independent experiments, each with three technical replicates and error bars represent the standard error of the mean for **(C)** and **(D)**. Unpaired t-tests were used to calculate statistical significance. **= $p<0.01$. CT= cycle threshold

Resources are more readily available for the detection of SARS-CoV-2 and no design of primers or antibody optimisation was needed. The published CDC N3 primer pair was selected for RT-qPCR detection as our colleague had previously tested the efficiency [226]. Wild-type and hACE2-NSK cells were infected by two variants of SARS-CoV-2, a European lineage closely related to the original Wuhan virus (isolate Edinburgh 2 and hereafter EDB2) and the more recently emerged Omicron variant (lineage BA.1). Although SARS-CoV-2 is known to replicate rapidly in human cell cultures, the growth kinetics in pig cells is not well-documented. To ensure we fully captured the growth rate, supernatants were harvested at 24, 48, 72 and 144 HPI. As expected, overexpression of hACE2 in NSK cells enhanced susceptibility to both variants of SARS-CoV-2. EDB2 can replicate in wild-type NSK but at a slower rate, indicated by virus titres peaking at 48 HPI in contrast to hACE2-NSK cells where titres peaked at 24 HPI (**Fig3.9A**). Interestingly, Omicron does not replicate in the absence of hACE2 in pig cells. Even in the presence of hACE2, the growth kinetics of Omicron is slower than EDB2, with only a small amount of virus detected at 48 HPI, although virus levels continued to increase throughout 72 and 144 HPI in hACE2-NSK (**Fig3.9B**). Based on these results, EDB2 was chosen as the preferred variant for this project, as it has faster growth kinetics and the ability to infect pig cells independently of hACE2 *in vitro*.

During the propagation of EDB2, it was found that the virus does not cause complete CPE on Vero cells, even with extended incubation. However, during initial testing of hACE2-NSK cells, extensive CPE was observed in cells infected by EDB2 starting at 48 HPI (**Fig3.9E and F**). We have therefore tested hACE2-NSK cells in place of Vero cells for titring SARS-CoV-2 infectious virus by TCID50. Supernatants of EDB2-infected wild-

type and hACE2-NSK cells were harvested at 24 HPI and titred on hACE2-NSK cells. The infectious virus titre of EDB2 in hACE2-NSK cells is 1 log higher than wild-type NSK cells at 24 HPI, consistent with the previous experiment where it is approximately 4 CT lower (**Fig3.9C**).

Similar to PRCV, the sensitivity of EDB2 towards universal and porcine IFN in hACE2-NSK cells was measured. Cells were treated with IFNs at various concentrations for 24 hours before being infected and the relative viral levels were determined by RT-qPCR from supernatant samples. The IC50 of both IFNs in hACE2-NSK cells were similar in restricting EDB2, ranging from 1 to 3 unit/ml (**Fig3.9D**). Compared with the result obtained from PRCV, the potency of universal IFN is similar in restricting PRCV and EDB2 in NSK-derived cells (**Fig3.10A**). However, interestingly, the potency of porcine IFN in restricting the two viruses is different. PRCV was less sensitive to porcine IFN than SARS-CoV-2, requiring roughly 10-fold higher concentration to achieve a 50% reduction in viral levels (**Fig3.10B**). This suggests, at least in the NSK cells, that PRCV is better adapted and able to evade restriction from porcine IFN treatment than the human SARS-CoV-2 virus.

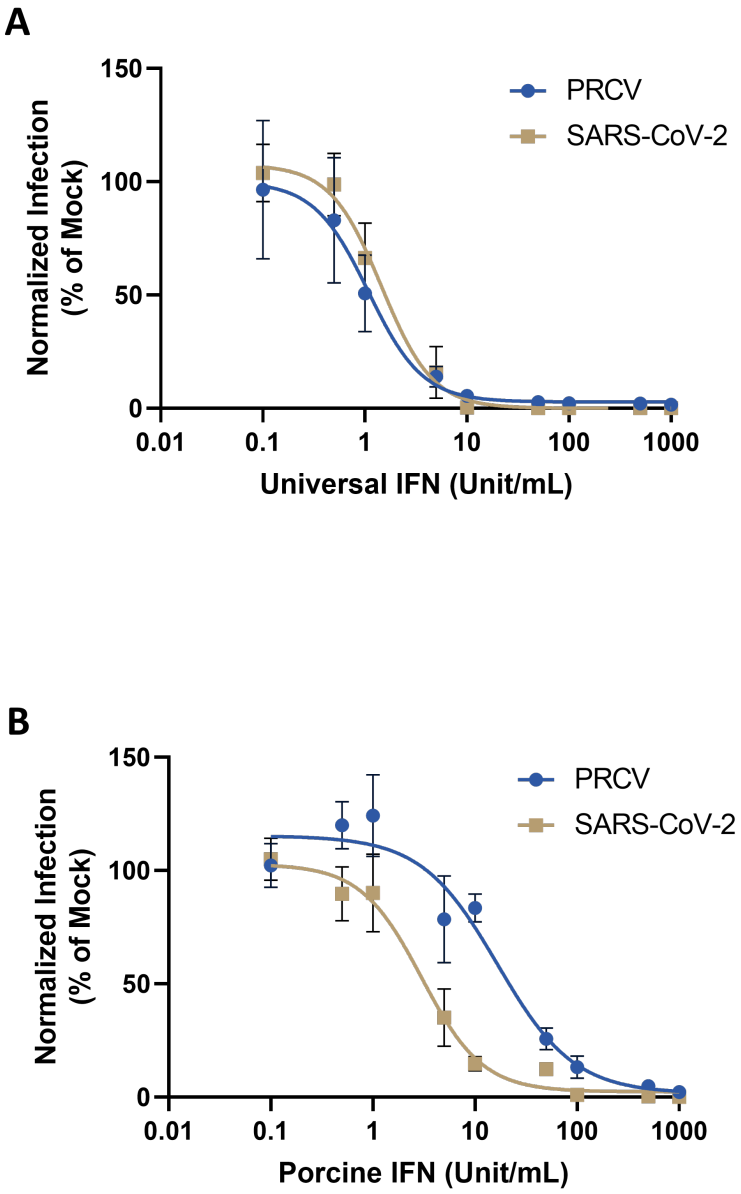


Figure 3.10: **Differential sensitivity of PRCV and SARS-CoV-2 towards universal and porcine IFN.** The sensitivity of PRCV and SARS-CoV-2 towards **(A)** universal Type I IFN and **(B)** porcine IFNα were compared. pAPN-NSK and hACE2-NSK were mock-treated or treated with either IFN at the indicated concentration for 24 hours before being infected by PRCV and SARS-CoV-2 isolate EDB2 respectively at an MOI of 0.01. Supernatant samples were harvested at 24 HPI and the relative infection level was determined by RT-qPCR and the 2^(-delta CT) method. Results represent the average of three independent experiments, each with three technical replicates and error bars represent the standard error of the mean. Data correspond to **Fig3.8D** and **Fig3.9D**.

3.3 Discussion

Identification of an accurate method for virus detection as well as selecting a suitable cell line for appropriate experiments were a vital step towards successful ISG screens in pig cells. In this chapter, we showed that the N gene of PRCV is an ideal target for RT-qPCR detection. As suggested in the first chapter, the S genes of coronaviruses have a relatively high sequence variability when compared to other genes and may not be an ideal target for RT-qPCR. The primer pair design was based on the only available reference genome at that time and mutations may have occurred during propagation, leading to unspecific binding. This may also have contributed to the attenuated phenotype observed in our initial PRCV stocks. While the primer pair targeting the ORF3 gene worked well, the primer pair targeting the N gene was selected as the efficiency was closer to 100% and matched the SARS-CoV-2 detection assay, which also targets the N gene.

The unique transcription mechanism of coronaviruses leads to the production of subgenomic mRNA as shown in **Fig1.1**. The N gene is located near the 3' end of all coronaviruses' genomes and although not translated, the sequence is present in most subgenomic mRNAs generated during replication. Therefore, the target sequences for primer pairs targeting the N gene are very abundant in coronavirus-infected cells, leading to high copy numbers in cell lysates. In contrast, each virion carries only one copy of the target sequence, explaining the lower abundance in supernatant samples. In cells that produce CPE, the subgenomic mRNA might be released, leading to an overestimation of the N copy number in the supernatant. Detecting viruses by RT-qPCR is indirect and techniques such as TCID50 and plaque assays remain the most direct approach for

quantifying infectious viruses. Moreover, unlike cell lysate samples, where housekeeping genes such as GAPDH and β-actin can be used for normalisation, the lack of appropriate loading controls for RT-qPCR of supernatant samples increases potential error. Nonetheless, RT-qPCR is less cumbersome, more rapid, and can be applied to a large number of samples simultaneously. Therefore, it remains a common approach for virus detection, especially in clinical settings.

Surprisingly, the only antibody that had shown reactivity towards PRCV with one of the tested approaches was a TGEV polyclonal antiserum. Two of the other three tested monoclonal antibodies reportedly target the S glycoprotein of TGEV, while the remaining have reactivity against the N protein of a wide range of veterinary coronaviruses, including FCoV, CCoV and bovine coronavirus. One possible explanation for the S-targeting antibodies failing to recognise PRCV is that they are targeting a region in TGEV that was deleted in PRCV. It is, however, surprising that the N-targeting antibody does not cross-react with PRCV. Nonetheless, we cannot dismiss the possibility of technical issues, as TGEV-infected samples were not included as positive controls. Despite this, we have successfully applied the antiserum for fluorescence microscopy, the concentration required (i.e. 1:1 undiluted), as well as its conjugated nature (i.e. signal amplification by secondary antibody not possible) has limited its use on a larger scale, such as high-throughput read-out for a screening experiment.

Although we have successfully enhanced the susceptibility of NSK cells towards PRCV by overexpression of entry receptors, attempts to generate susceptible human cell lines using the same approach were unsuccessful. Initially, it was thought to be due to the

IFN response in the primary fibroblast cells and the attenuated phenotype of PRCV. However, subsequent experiments suggest expressing pAPN in human cells with IFN-related deficiency is insufficient to induce susceptibility to PRCV or TGEV, despite published reports to the contrary. In the study carried out by Li *et al.*, the authors also HA-tagged pAPN at the C-terminal, showing that transfection and transduction rendered HeLa cells susceptible to TGEV infection and rescued the phenotype of pAPN knockout ST cells [252]. Wang *et al.* have also shown that Vero and BHK-21 cells stably expressing C-terminal Myc-tagged pAPN produce infectious TGEV by TCID50 [251]. By analysing the methods of these studies, it was realised that a virulent Purdue strain, in contrast to a TC-adapted Miller strain, was used in the published experiments. It remains possible that the attenuated TGEV and PRCV we used lack certain factors that are required to infect non-porcine cell lines. Although the entry of PRCV into human HEK293T cells expressing pAPN has been reported, the experiment was based on lentiviral pseudotypes expressing PRCV S protein [254]. Nonetheless, the result of this study suggests the blockade likely occurs at a post-entry step, with the IFN response as one of the potential candidates. Repeating the pseudotype experiment in more cell lines, along with the use of different inhibitors such as ruxolitinib (a JAK inhibitor), can help determine whether it is possible to render human cell lines truly susceptible to PRCV.

The relatively high traduction efficiency of NSK cells prompted us to test their susceptibility to SARS-CoV-2, as this will allow us to perform the screening experiment against the two viruses in the same cell background for direct comparative analysis. Using an early SARS-CoV-2 isolate EDB2 and a later variant Omicron, we show that

EDB2, but not Omicron, can replicate in wild-type NSK cells. Through overexpression, hACE2-NSK cells were able to support EDB2 replication to an even larger extent and become susceptible to Omicron. The data suggests that EDB2 can utilize pACE2 for entry while Omicron is highly dependent on hACE2 *in vitro*. The amino acid sequence similarity between hACE2 and pACE2 is around 80%. However, the affinity of EDB2 towards hACE2 and pACE2 was not compared in this study. Whether the higher EDB2 titre observed in hACE2-NSK was due to the specific expression of hACE2 or simply the overall increased expression of entry receptors (irrespective of human or porcine origin) is yet to be determined.

The Omicron variant was first reported in late 2021 after it had been circulating in the human population for nearly two years. Multiple mutations have been reported in the RBD of S protein and have been suggested to allow Omicron to become better adapted to human receptor usage than earlier variants. Indeed, multiple studies have reported that the S protein of Omicron has a higher affinity to hACE2 than earlier variants [255–258]. Although Omicron is not tested, a study comparing the affinity of S protein from Alpha, Beta, Gamma and Delta variants did show that they have differential affinities towards ACE2 from different animals [259]. It is possible that less adapted variants, such as the early EDB2 can utilize ACE2 receptors for entry from a broader range of species, including pigs, while Omicron is more specialised in the use of hACE2 for entry.

Studies have suggested, that SARS-CoV-2 variants of concern (VOC) including Alpha, Beta, Gamma, and Delta arise during adaptation in the human population had evolved convergent strategies to suppress ISGs in humans [260, 261]. In contrast, early Omicron

variants (i.e. BA.1, the lineage used in this project) did not, possibly related to its origination from a chronically ill and immunocompromised patient [262, 263]. Similar to receptor usage, it is possible that EDB2 can suppress the innate immune response of a broader range of species, therefore able to infect the WT NSK cells, while early Omicron variants require the overexpression of entry receptors (i.e. increased viral dosage) to compensate for its weaker suppression potency.

The entry of 'pre-Omicron' variants is reported to be more dependent on the presence of transmembrane protease serine 2 (TMPRSS2), which cleaves the S2 domain and allows fusion on the cell surface. In contrast, Omicron prefers entry by the endocytic route, where fusion occurs in the acidic endosome in the presence of cathepsin [264, 265]. Interestingly, Omicron entry into cells with high TMPRSS2 levels was reported to be less efficient compared to other variants [266]. As kidney cells are known to express high levels of TMPRSS2, it is possible that wild-type NSK cells (a kidney cell line) allow efficient entry of EDB2 but impair or limit the entry of Omicron. However, overexpression of hACE2 may have overcome this inhibition as shown by the data from hACE2-NSK cells.

Prior to further experiments, the responsiveness of the selected cell lines towards universal and porcine Type I IFNs was determined, allowing us to confirm the IFN competence of NSK-derived cells. The sensitivity of PRCV and SARS-CoV-2 towards the two IFNs was also compared, providing evidence that IFNs can have differential potency against coronaviruses of a different animal origin. Both universal and porcine IFN have a similar potency against SARS-CoV-2 in NSK cells. However, they only have a similar potency

against PRCV in ST cells, as a higher concentration of porcine IFN is required to achieve similar restriction in NSK cells. The data suggests virus sensitivity to IFN is IFN-type and cell-type specific. Despite all binding to the same IFNAR1/2, numerous reports have suggested various Type I IFNs can trigger a different response by differential ligand-receptor interactions [267]. Although the universal IFN has a human IFNα backbone, it has been mutated to maximise its potency in various mammals, possibly making it behave differently from canonical IFNα. Indeed, as mentioned above, the responses triggered by the universal and porcine IFN in primary pig cells are different [154]. The responses triggered by IFN are also known to vary across cell types due to multiple differences, such as underlying epigenetic features [268]. It is therefore not surprising that ST and NSK cells may have different sets of up- and down-regulated genes following the same IFN treatment.

Consistent with our hypothesis, the porcine-adapted virus PRCV is less sensitive to porcine IFNα response than the non-adapted virus SARS-CoV-2 in NSK-derived cells. This suggests PRCV can overcome or evade certain ISGs in NSK cells which SARS-CoV-2 is more sensitive to. The data provided strong grounds for ISG overexpression screening experiments against the two viruses in the generated pAPN- and hACE2-NSK cells, allowing us to interrogate the function of each ISG independently and potentially pinpoint a specific ISG that contributes to the differential sensitivity observed.

CHAPTER 4

Identification of porcine coronavirus restriction genes using arrayed lentiviral ISG Screening

4.1 Overview and Background

Viruses must overcome and adapt to barriers such as the IFN/ISG response posed by new hosts to enable zoonotic transmissions. In the previous chapter, we have shown that the porcine non-adapted SARS-CoV-2 is more sensitive to porcine IFN than the adapted PRCV. To identify key components of the porcine IFN response responsible for this differential sensitivity, we adopted an ISG overexpression screening approach, which has proven effective in dissecting the human IFN response.

According to the latest human genome assembly GRCh38, more than 20,000 protein-coding genes were identified. Viruses have the potential to interact with a wide range of proteins and the traditional way of understanding virus-host interaction by interrogating one gene at a time is time-consuming and may lack appreciation of context in relation to multiple complex interactions occurring in the infected cell. With the advances in technology, screening experiments where hundreds and sometimes thousands of genes can be studied simultaneously have become a powerful tool to identify key host proteins involved in promoting or restricting virus replication.

The most widely used approach for defining a gene function is to knock down or knockout its expression (i.e. loss-of-function). The use of RNA interference (RNAi) and the more recent CRISPR/Cas9 systems are popular tools for loss-of-function screens. In brief, RNAi targets and cleaves mRNA in the cytoplasm, achieving knockdown at the transcriptome level; while Cas9 coupled with guide RNAs (gRNAs) target genes and create double-strand breaks, leading to mutation and achieving knockout at the

genomic level. These approaches have been successful in identifying numerous host factors involved in many human viral infections, including HIV, HCV, IAV, Zika virus and of course, SARS-CoV-2 and other HCoVs [269–276].

However, loss-of-function screens may not be as effective at identifying factors restricting replication, especially those that have a modest effect. One reason is redundancy, so reducing the expression of one restriction factor may not necessarily produce a detectable phenotype. It is also easier to measure reduced virus production than an increase due to differences in dynamic range, as the loss of an essential host gene can lead to a complete restriction of virus replication, which in most cases will be multiple logs. Whereas knockout of a restriction factor will rarely lead to more than a ten-fold increase as there is a limit to how much virus can be produced by a cell. Gain-of-function or overexpression screens have therefore become an alternative approach. Nonetheless, overexpression also has its drawbacks, including toxicity, inability to study proteins that function in a complex, and less relevance to actual physiological conditions [277]. It is therefore common for studies to utilise both approaches to compensate for each drawback.

The recent development of a catalytically dead version of Cas9 (dCas9) has enabled the adaptation of CRISPR technology to overexpression screens. CRISPR activation (CRISPRa) utilizes dCas9 fused to a transcription activator. Instead of cleaving the targeted gene, these fusion dCas9s combine with gRNA and bind to target genomic locations (usually promoter region), allowing the activation of selected genes [278–280]. CRISPRa has several potential advantages over traditional techniques, such as the ability to generate all gene isoforms and enable us to study gene regulation in *cis*. Regardless,

this is a relatively new technique with limited reports, especially for studying virus-host interaction [281]. Most overexpression screens are therefore achieved by cloning the cDNA of protein-coding genes into expression plasmids, followed by direct transfection or transduction with lentiviruses in selected cells. Genome-scale lentiviral overexpression libraries have been used to understand drug mode-of-action and resistance mechanisms as well as melanoma development in the past [282, 283]. However, some of the most successful studies used overexpression screens to identify antiviral ISGs.

As mentioned, the human genome is relatively large. A genome-scale library can therefore consist of tens of thousands of plasmid cDNA. This number can be even larger in CRISPR and siRNA libraries, as multiple gRNA or siRNA were designed for each gene. Genome-wide screens are therefore labour-intensive when performed in an arrayed format without the help of robotics and may require more optimisation. Moreover, some studies may only be interested in a specific group of genes and the use of genome-wide libraries is unnecessary. Targeted libraries comprising a smaller subset of genes have therefore been developed. Examples include the druggable genome library, a DNA binding domain-focused library and in the case of this project, ISG libraries.

The definition of an ISG is a gene that is upregulated following the IFN response. The most common approach to identify ISGs is RNA-seq which allows the quantification of RNA. By comparing the transcriptome of IFN-treated and untreated cells, specific transcripts upregulated following treatment can be identified. However, the degree of upregulation among ISGs can vary substantially, ranging from 1.1 to 1000-fold increase. There is no universal consensus on the level of upregulation required by a gene to be

defined as an ISG, with studies setting cut-offs ranging from 1.5-fold to 4-fold [109]. Alternatively, instead of using a biological cut-off, some studies use a statistical cut-off, defining ISG based on false discovery rate (FDR) or p-value regardless of the extent of upregulation [154].

In 2011, Schoggins *et al.* developed the first lentiviral overexpression library dedicated to ISGs. The library, consisting of 389 genes, was constructed based on ten publications with microarray data sets from various type I IFN-treated cell types or tissues. In brief, the ISGs were cloned into a lentiviral vector to generate lentiviruses, which were used to transduce target cells, allowing the overexpression of individual ISGs and a marker RFP (similar to pSCRPSY, see **3.2.1** for details). Transduced cells were then challenged with a GFP-expressing virus, and viral replication within the RFP-positive population was quantified by fluorescence-activated cell sorting (FACS). The library was applied to six different viruses, identifying a diverse range of both broadly acting and specific antiviral effectors [109]. The initial screen also identified the board antiviral potency of the DNA sensor cGAS (previously known as *C6orf150*) before its molecular mechanism was resolved [284, 285]. Since then, this library has been expanded to include more ISGs and separate libraries were developed for type II and III ISGs [249, 286, 287]. ISG libraries have been applied not just to viruses, but also bacteria and parasites, identifying various key factors in restricting a broad range of pathogens [111, 221, 250, 287–290].

The use of such ISG libraries has been valuable in allowing us to understand how different human ISGs restrict various human pathogens. However, a smaller subset of ISGs is repeatedly identified, as more screens are being reported, suggesting most

human pathogens may be restricted by a smaller subset of ISGs. Indeed, a recent study demonstrated that IFN-mediated suppression of viruses is dominated by a few ISGs out of hundreds of IFN-inducible mRNAs [291]. To evaluate the potential effect of other ISGs, the trend of applying these libraries has switched to animal pathogens to identify potential species-specific restriction factors that human-adapted viruses can evade, but are potently antiviral against non-adapted animal viruses. Such species-specific restriction factors may play important roles in restricting zoonotic disease. Indeed, Pinto *et al.* identified a primate-specific ISG that potently inhibited avian-adapted IAVs but had little impact on human-adapted IAVs [181].

Screening resources including CRISPR, siRNA and lentiviral overexpression libraries in humans are well-established, allowing us to uncover the role of different genes during virus infections. However, libraries from other animals, apart from model organisms such as monkeys and mice, are lacking. With the advances in sequencing techniques, the annotation of genomes of other animals has improved, allowing us to generate resources and perform screening experiments. Livestock species such as pigs and chickens are of great interest, as they have high economic values and can act as intermediate hosts for zoonotic diseases. Our group has therefore developed genome-wide CRISPR/Cas9 and type I ISG libraries for these two species.

A cut-off of log2 fold-change > 2 and FDR < 0.05 were applied to the RNA-seq data from the study of Shaw *et al.* to comprise an initial list of porcine ISG [154]. Based on the porcine genome assembly Sscrofa10.2, poorly defined genes were removed, while some known antiviral ISG from human studies were added to the list to serve as positive

controls [221]. In this project, a final library comprising 432 pig ISGs will be applied in the porcine cell lines generated in the previous chapter, aiming to identify porcine ISGs that have an antiviral effect on coronaviruses.

4.2 Results

4.2.1 Optimisation of arrayed lentiviral ISG screens

ISG screens in human cells have been widely reported and are relatively well-optimized. However, the performance of these screens in cells of different animal origins was relatively novel at the time this project started. Nonetheless, optimisation was carried out based on the experience of previous screens performed by our group in human NHDF cells [250].

In the previous chapter, we have shown that the porcine-adapted PRCV is less susceptible to porcine IFN than the non-adapted SARS-CoV-2 in NSK-devired cells and ISG screening can potentially identify cross-species restriction factors. We have also shown that NSK cells can be modified to become highly susceptible to the two viruses while being transducible with the selected lentiviral vectors. However, prior to performing a medium-to-high throughput experiment, several aspects such as dynamic range, read-outs, transduction level after scaling down and sample storage have to be optimised.

All screening optimisations were performed using PRCV instead of SARS-CoV-2 as the latter requires a biosafety level 3 facility with limited access. Before further optimisation, a potent positive control ISG must be identified as a positive control, enabling the determination of the maximum dynamic range. Components of the IFN signalling pathways such as the PRRs, IRFs and STATs are known ISGs and overexpression should restrict virus replication by triggering an IFN response. Lentiviruses encoding porcine IRF1, IRF7, STAT2 and MDA5, as well as a GFP control were generated and used to

transduce ST, NPTr and pAPN-NSK cells. It was shown in the previous chapter(**Fig3.6**) that, despite the three cell lines having different transduction efficiencies (represented by RFP signal intensity), their ability to express the inserted ISG (represented by GFP signal) is similar. However, surprisingly, none of the NPTr cells transduced with the selected ISGs showed a reduction in PRCV replication. In contrast, all transduced ST cells, including the negative control GFP reduced PRCV replication. Cytotoxicity was also observed in ST cells suggesting that, despite being more transducible, they are also more sensitive to the cytotoxic effect from transduction (data not shown).

The pAPN-NSK cells transduced with IRF1 and IRF7 significantly reduced PRCV N copy number (**Fig4.1A and B**). IRF7 was more potent, reducing virus level by 3 logs and 1.5 logs in cell lysate and supernatant samples respectively, compared to IRF1, where virus replication was reduced by 1 log in both samples. Cells transduced with STAT2 and MDA5 did not show any sign of reduction in viral signal. Based on these results pAPN-NSK cells were selected for further experiments and IRF1 and IRF7 as positive controls.

One of the biggest challenges in optimising high-throughput screening experiments is defining a robust read-out method as this will directly affect the accuracy of the results. The majority of previous ISG screens have applied reporter viruses expressing GFP or luciferase, and the viral level can be directly determined by signal intensity to identify hits [105, 249, 250, 288, 290]. Although we have access to a GFP reporter SARS-CoV-2, a PRCV GFP reporter virus is not currently available. The use of reporter viruses is therefore not possible for this project. An alternative is the use of antibodies that

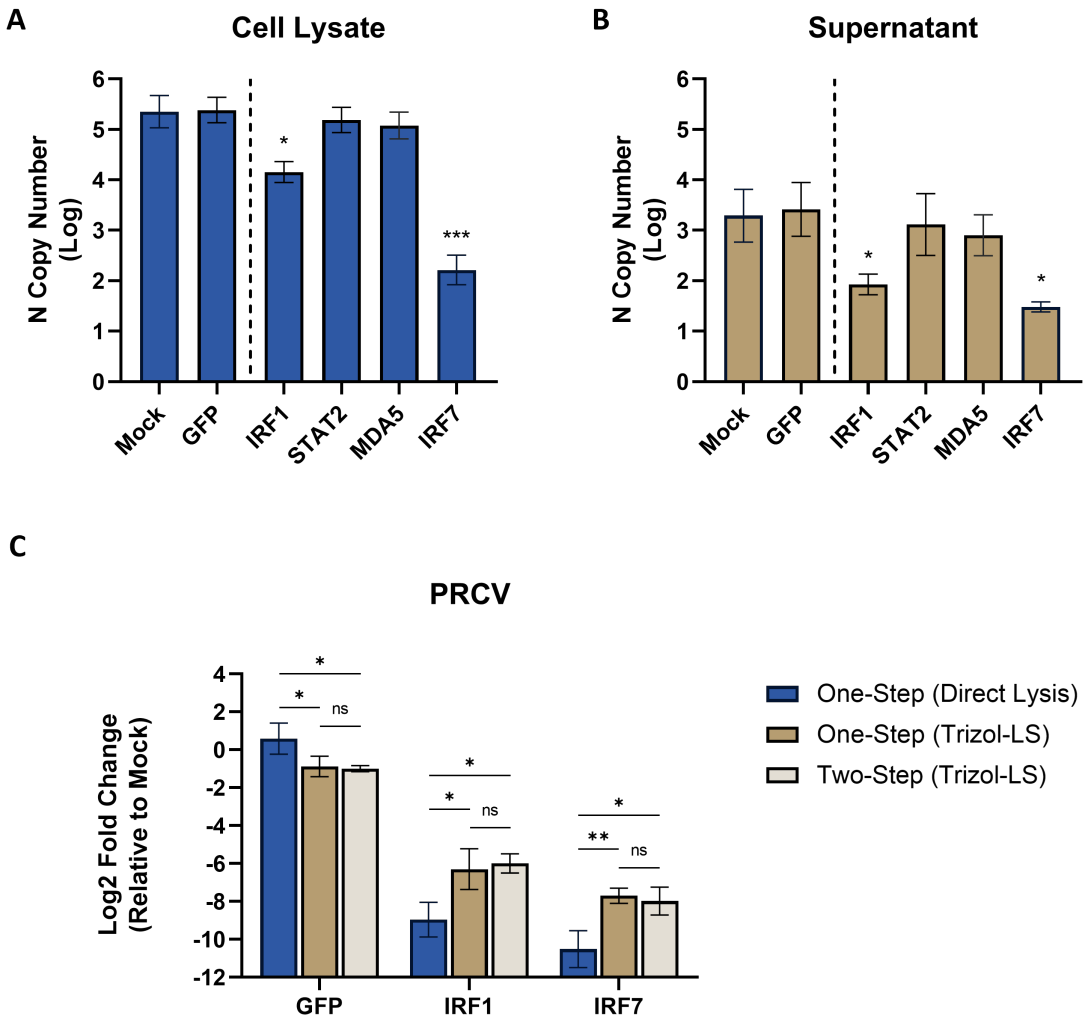


Figure 4.1: **IRF1 and IRF7 as positive control and read-out optimisation**. pAPN-NSK cells were seeded in 24-well plates and transduced by lentivirus expressing the indicated genes. Transduced cells were infected by PRCV at an MOI of 0.01 at 3 days post-transduction and samples were harvested at 24 hours post-infection. The N copy number of **(A)** cell lysate and **(B)** supernatant samples was calculated using a standard curve generated by a control plasmid. **(C)** Supernatant samples from separate experiments were either directly lysed and used as input for one-step RT-qPCR or subjected to conventional Trizol-LS RNA extraction. Extracted RNA was either used as input for one-step RT-qPCR or subjected to conventional reverse transcription with the resulting cDNA used as input for two-step RT-qPCR. The PRCV levels were reported as Log2 fold change relative to mock transduced cells. Results represent the average of three independent experiments, each with two technical replicates. Error bars represent the standard deviation. Multiple comparisons were performed from a one-way ANOVA, comparisons were against the control GFP for **(A)** and **(B)**. *=$p<0.05$, **=$p<0.01$, ***=$p<0.001$, ns= not significant.

recognize viral proteins [221]. However, although antibodies were well developed for SARS-CoV-2 detection, antibodies for PRCV detection are sub-optimal as discussed in the previous chapter. Another method is therefore needed for consistent read-outs for both viruses.

RT-qPCR has been shown to be a reliable method for both PRCV and SARS-CoV-2 detection in experiments from above and the previous chapter, although from larger 24-well plates, with relatively large volumes of sample available for Trizol-LS RNA extraction. However, the screening experiments would be performed in 96-well plates, and RNA extraction from small volumes of samples was shown to be difficult. We have been detecting both cellular and supernatant viral RNA but the latter is a better surrogate of produced viral particles. Rather than extracting RNA, our colleagues developed a virus lysing buffer (VL buffer) for direct measurement of SARS-CoV-2 from supernatant samples [226]. This protocol was therefore modified for PRCV detection.

To compare results obtained by VL buffer and Trizol-LS RNA extraction, cells in 24-well plates were transduced by GFP, IRF1 or IRF7 lentivirus and infected with PRCV. 5µl of supernatant were directly lysed using VL buffer with total RNA extracted from the remaining supernatant using Trizol-LS. The VL buffer lysate was diluted 10-fold before being used as input for RT-qPCR (One-Step, Direct Lysis). 2µg of the extracted RNA was subjected to conventional reverse transcription and 200ng of cDNA was used as input for qPCR (Two-Step, Trizol-LS), while 200ng of RNA was directly used as input for RT-qPCR (One-Step, Trizol-LS) (**Fig4.1C**). The results showed no statistical difference between one-step and two-step RT-qPCR of extracted RNA and cDNA. Although the

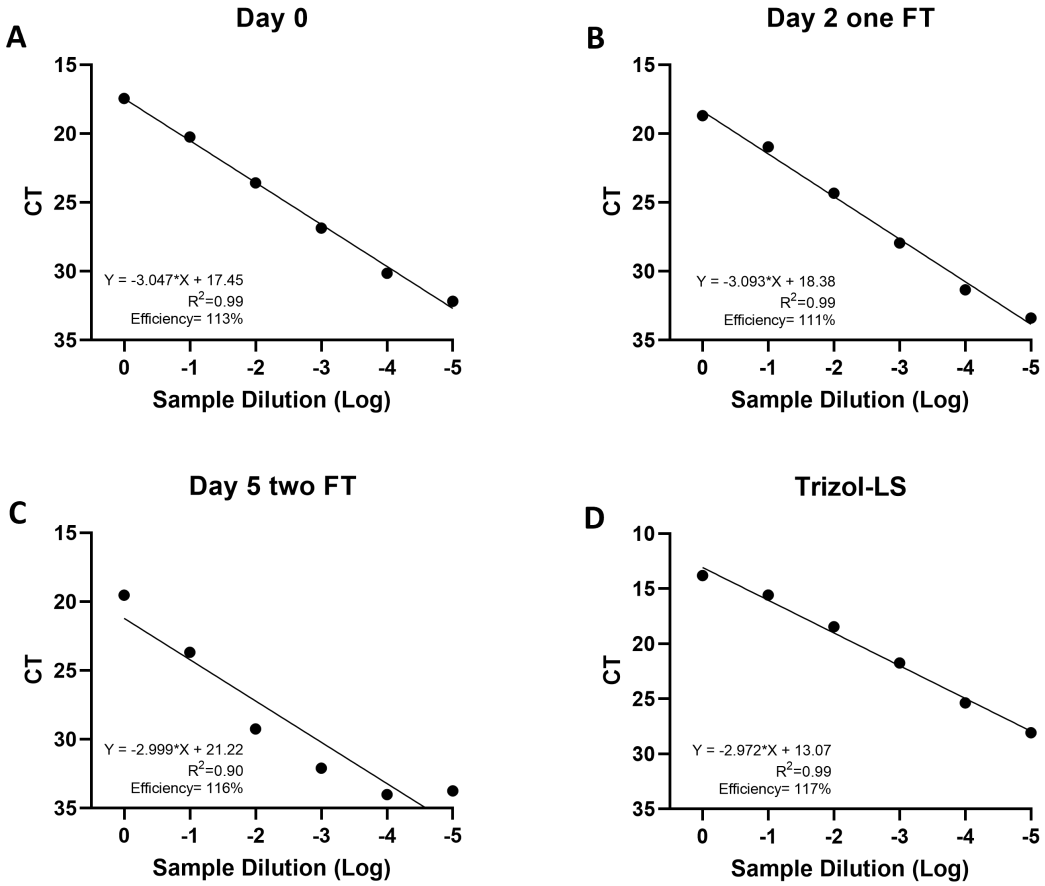


Figure 4.2: **Stability of lysate sample undergoing freeze-thaw cycle.** An aliquot of PRCV stock was serially diluted before being lysed or RNA extracted and stored at -20°C. The samples were subjected to one-step RT-qPCR on **(A)** day 0 (no freeze-thaw), **(B)** day 2 (one freeze-thaw) and **(C)** day 5 (two freeze-thaw). **(D)** The extracted RNA was subjected to one-step RT-qPCR on day 5 (one freeze-thaw) as control. Linear regressions were calculated to assess reaction efficiency from the slope. Results represent the average of three technical replicates. Error bars were not shown due to the small standard deviation from the replicates. CT= cycle threshold, FT=freeze-thaw.

results obtained by direct lysis are statistically different, a similar trend was observed, with IRF1 and IRF7 both restricting PRCV potently. This suggests the direct lysis protocol is reliable for relative PRCV quantification and can be applied as the read-out for screening experiments.

The number of samples in pilot experiments is limited and can be easily processed within a short period of time. However, during the screening experiment, a large number of samples will be produced and it is not practical to process them within the same day. The samples would need to be stored properly in an appropriate condition before use. To test the stability of lysed samples from the VL buffer, an aliquot of PRCV stock was serially diluted before being lysed or RNA extracted and stored at -20°C. The samples were subjected to one-step RT-qPCR on day 0 (no freeze-thaw), day 2 (one freeze-thaw) and day 5 (two freeze-thaw) (**Fig4.2**). There was little difference between the samples processed on day 0 and day 2, suggesting the lysate remains stable at -20°C for at least 2 days and one freeze-thaw cycle. However, two freeze-thaw cycles resulted in reduced signal and greater variation, especially in lower copy number samples. As expected, RT-qPCR on extracted RNA was more sensitive and produced a lower CT ($\sim$ 4 CTs) than lysates from the VL buffer at the same dilution factor (**Fig4.2D**).

When identifying positive controls in **Fig4.1**, it was to our surprise that STAT2 and MDA5 did not reduce PRCV titre, despite being potent hits, likely by being potent IFN inducers, in other screens [109]. We speculate this might be due to insufficient transduction levels as the experiments were performed in 24-well plates (i.e. a lower virus-to-cell ratio). Some reports have suggested spinoculation can enhance lentiviral

transduction efficiency in certain cell lines. However, in pilot experiments, it was found that this approach cannot be applied to 24-well plates due to the size of the wells, as the centrifugation force pulls the medium towards the edges, drying the cell monolayer in the process. To access the effect of spinoculation and characterize the relationship between transduction level and phenotype, pAPN-NSK cells were seeded in 96-well plates and transduced with a panel of defined antiviral ISGs by three different methods: forward transduction, forward spinoculation and reverse spinoculation.

Forward transduction refers to the conventional method used in all previous experiments; forward spinoculation is similar to forward transduction, with an additional step of centrifuging the plates at 800 x *g* for 10 minutes after lentivirus addition; reverse spinoculation is similar to forward spinoculation, except cells were seeded as lentivirus was added. Following infection, the PRCV level was determined using direct lysis. Consistent with **Fig4.1**, only IRF1 and IRF7 showed significant decreases in PRCV levels when cells were forward transduced. Spinoculation substantially enhanced transduction efficiency as shown by an increased RFP fluorescence intensity (**Fig4.3A, C and E**). The RFP signals from certain ISGs such as MDA5 and IRF7 remain similar across the three methods. Nonetheless, MDA5 showed a significant decrease in PRCV levels only through spinoculation, irrespective of fluorescence intensity. Otherwise, most ISGs showed a dose-dependent effect, exhibiting higher PRCV restriction with higher transduction efficiency (**Fig4.3B, D and F**). However, a large decrease in PRCV levels (-4 Log2) in cells transduced with the negative control GFP following reverse spinoculation suggests that excessive transduction can lead to unspecific inhibition, most likely due to the IFN response triggered by increased amounts of lentivirus. Indeed, cells that have undergone

reverse spinoculation showed signs of cytotoxicity and abnormal morphology (data not shown), suggesting this is not an ideal approach to enhance transduction efficiency for NSK cells. Although a significant difference is also observed in cells forward spinoculated by GFP, the biological effect is relatively minor (-1.5 Log2) and can be corrected by normalization (**Fig4.3D**).

The fluorescence intensity of cells reverse spinoculated with GFP lentivirus has exceeded the detection limit and the level of transduction at which unspecific inhibition starts to occur cannot be determined. To address this, pAPN-NSK cells were forward spinoculated with a 2-fold dilution of GFP lentivirus and PRCV level determined by direct lysis RT-qPCR. The result showed that no significant differences in PRCV levels were found in GFP-transduced cells with fluorescence intensities below 50,000 (from 4- to 32-fold diluted). In contrast, cells with intensities above 60,000 (neat and 2-fold diluted) exhibit unspecific inhibition, significantly reducing PRCV by approximately 4-fold (**Fig4.3G and H**). All experiments involving transduction described hereafter were performed in 96-well plates using forward spinoculation with transduction levels adjusted to be no higher than 50,000 fluorescence units to avoid unspecific inhibition.

All lentiviruses used in pilot experiments were generated from T75 culturing flasks. However, this is not practical when producing lentivirus from the ISG library comprising 432 plasmids. The production has to be scaled down to 96-well plates and the titre and purity of produced lentivirus is inevitably reduced. The protocols for generating lentivirus from the two formats are different (see Chapter 2 for details) and optimisation was required to ensure the titre from 96-well plates is above the required threshold to

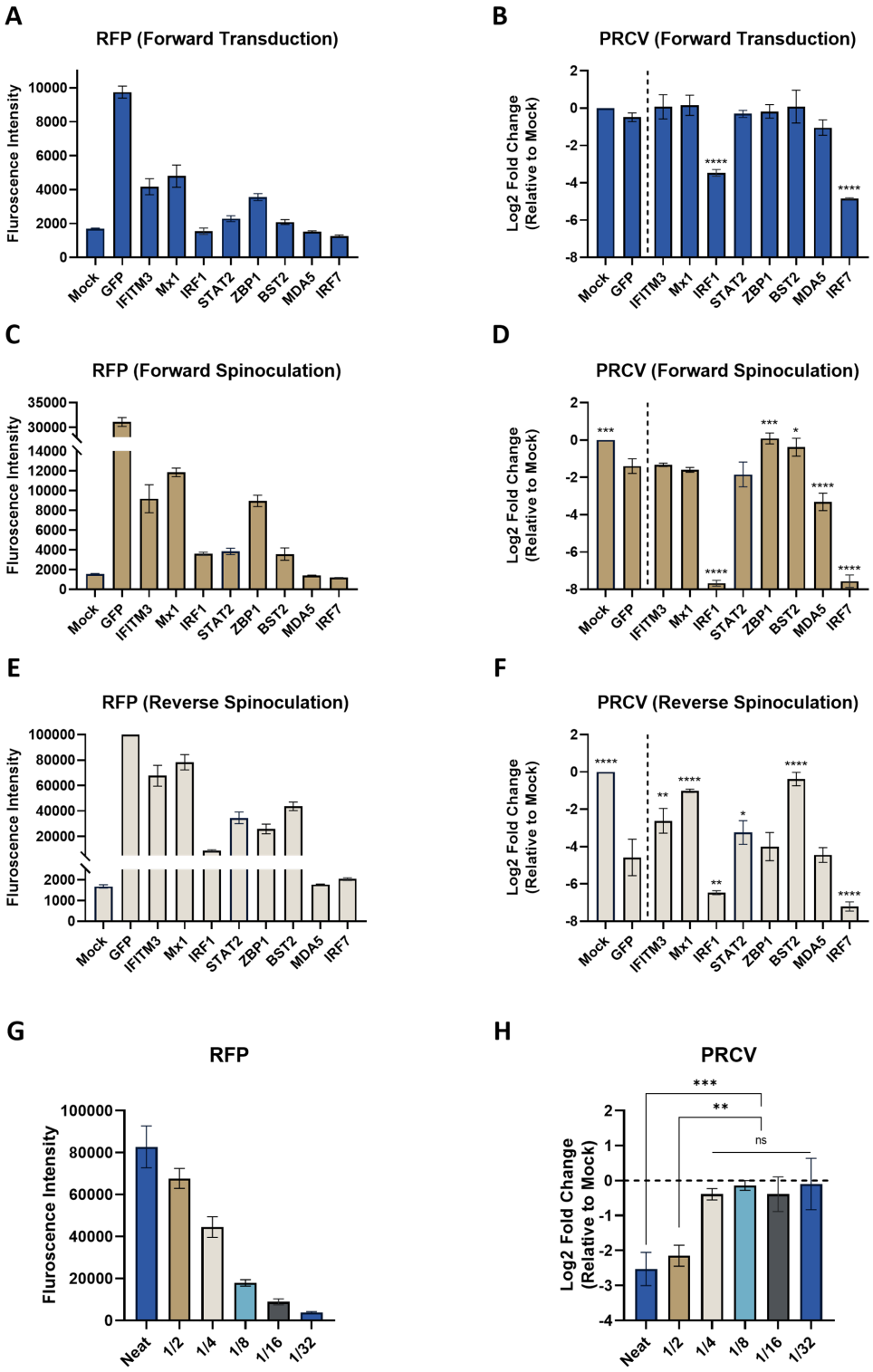


Figure 4.3: **Excessive transduction leads to unspecific inhibition.**

Figure 4.3: **Excessive transduction leads to unspecific inhibition.** pAPN-NSK cells were seeded in 96-well plates and transduced with indicated lentiviruses. RFP signal intensity was monitored daily and the readings on day 3 were reported for **(A)** forward transduction, **(C)** forward spinoculation and **(E)** reverse spinoculation. The signal from GFP in **(E)** has exceeded the detection limit. Transduced cells were infected by PRCV at 3 days post-transduction. Supernatant samples were lysed and processed by RT-qPCR at 24 HPI. The Log2 fold changes in PRCV levels relative to mock transduced cells were reported for **(B)** forward transduction, **(D)** forward spinoculation and **(F)** reverse spinoculation. An aliquot of GFP lentivirus was serially 2-fold diluted and used to transduce pAPN-NSK cells by forward spinoculation. **(G)** RFP signal intensity on day 3 and **(H)** PRCV levels were reported. Results represent the average of three independent experiments, each with two technical replicates. Error bars represent the standard deviation. Multiple comparisons were performed from a one-way ANOVA, comparisons were against the control GFP for **(B)**, **(D)** and **(F)**.****= $p<0.0001$, ***= $p<0.001$ **= $p<0.01$, *= $p<0.05$, ns= not significant.

produce a phenotypic effect. Lentivirus is reported to be unstable at room temperature or 4°C and freeze-thawing is known to reduce virus titre. Storage conditions for harvested lentivirus needed to be determined. To compare which condition can minimise the reduction in titre, an aliquot of GFP lentivirus was thawed and split in two, one kept at 4°C and the other returned to the freezer to undergo an extra round of freeze-thawing. After 6 days, the lentivirus was used to transduce pAPN-NSK cells and the RFP signal intensity was recorded. The result shows little difference in lentiviral titre was observed between the two conditions, and the harvested supernatant can be temporarily stored in either condition (**Fig4.4A**).

A GFP lentivirus was also produced during the porcine ISG library production in 96-well plate format for titration. To determine the approximate titre of the production batch, an aliquot of GFP viruses produced in T75 culturing flasks was 2-fold serially diluted and used to transduce pAPN-NSK cells in 96-well plates alongside the GFP lentivirus produced in 96-well plates. The result suggests the titre of lentivirus produced in 96-well plates is between 1/4 and 1/8 of that produced in T75 flasks (**Fig4.4B**). Combining this

with the data in **Fig4.3G and H**, where only GFP lentivirus with a concentration above 1/4 triggered unspecific inhibition, this suggests the majority of lentiviruses produced in 96-well plates were at an appropriate titre for transduction.

As stated in the previous chapter, lentivirus vectors have a limited packaging capacity of approximately 10 kb. Packaging efficiency will drop significantly once it reaches this limit. The size of the ISG inserts in the library varies from 200 bp to 5000 bp. Combined with the Puro-2A-TagRFP cassette and other essential elements for packaging, some constructs approach this limit and may have a lower titre due to packaging efficiency. To determine the correlation between size and transduction efficiency, the RFP fluorescence intensities were recorded from pAPN-NSK cells transduced with the library and plotted against their respective size. The coefficient of determination (R^2) between the two variables is 0.25, suggesting there is a weak correlation (**Fig4.4C**). Nonetheless, these ISGs only comprise a small portion of the library and did not affect the overall quality of the screen.

With optimisation completed, a finalised workflow for the screening experiments was established (**Fig4.5**). Briefly, following the generation of the lentivirus ISG library from packaging x293T cells in 96-well plate format, target cells seeded in 96-well plates would be transduced by the library through forward spinoculation. Three days post-transduction, the cells would be infected with the virus of interest and the supernatant harvested at 24 HPI. Harvested samples would be lysed using VL buffer and stored at -80°C until subjected to RT-qPCR for the determination of relative virus levels.

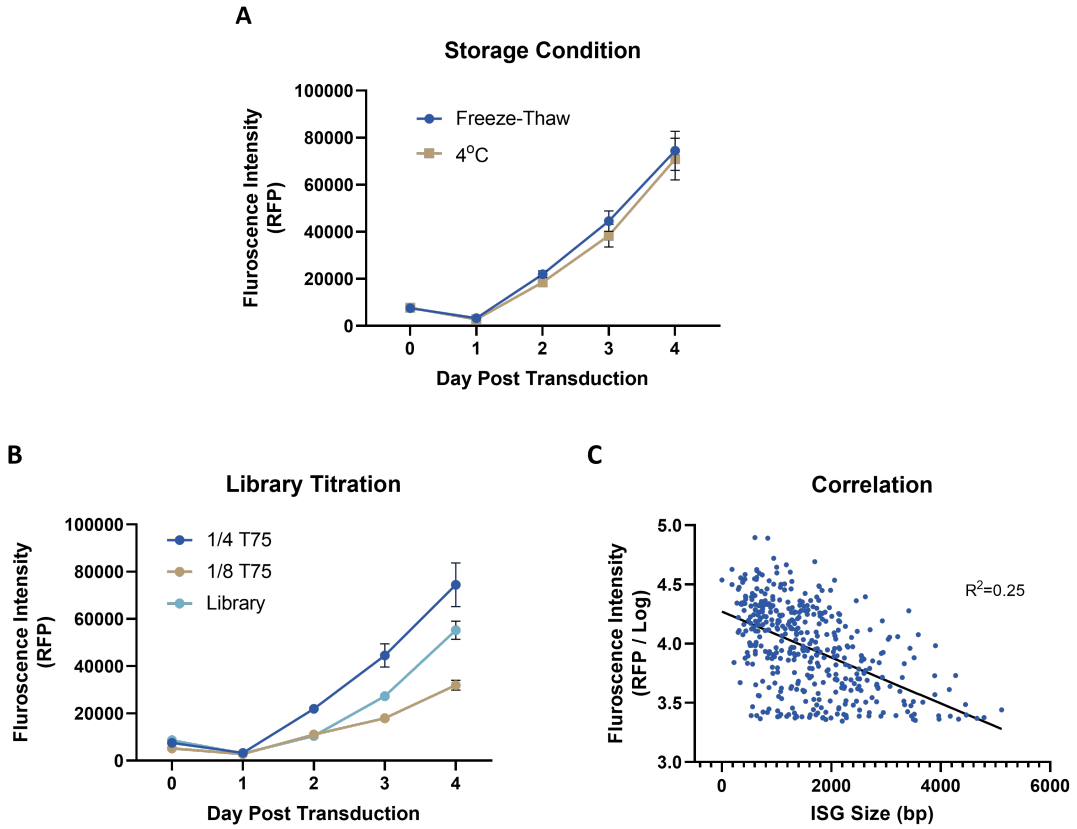


Figure 4.4: **Porcine ISG library production and characterisation**. **(A)** An aliquot of GFP lentivirus was thawed and either kept at 4°C or returned to the freezer for an extra round of freeze-thawing. The lentivirus was used to transduce pAPN-NSK cells in 96-well plates by forward spinoculation after 6 days. RFP fluorescence intensities were recorded daily and reported. **(B)** A GFP lentivirus was produced with the porcine ISG library in 96-well plate format. The titre of the library was determined by transducing pAPN-NSK cells with GFP lentivirus alongside the 2-fold serially diluted GFP lentivirus produced from T75 culturing flasks. **(C)** pAPN-NSK cells in 96-well plates were transduced with the ISG library. RFP fluorescence intensities were recorded daily, and readings from day 3 were reported and plotted against the size of the ISG inserts. The coefficient of determination (R^2) was calculated from a simple linear regression. Results represent the average of three technical replicates. Error bars represent the standard deviation. bp=base pair.

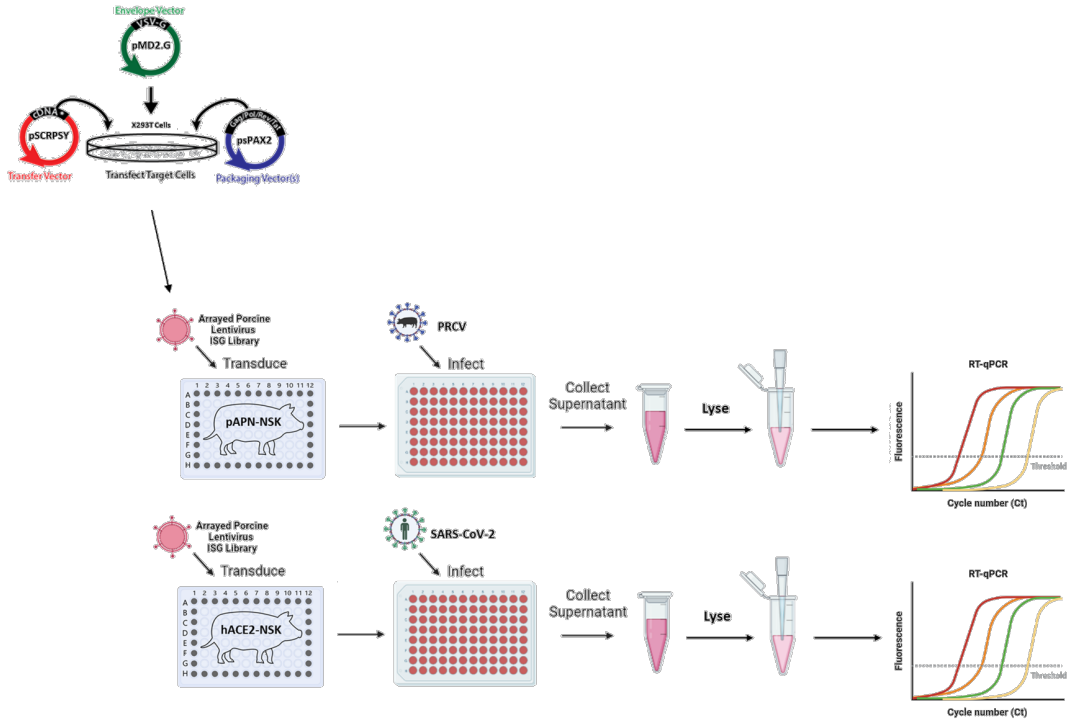


Figure 4.5: **Finalised workflow for arrayed lentiviral screening experiment**. The lentiviral library comprising 432 porcine ISGs was generated from packaging x293T cells in 96-well plate format. pAPN-NSK or hACE-NSK cells were seeded in 96-well plates and transduced with the library through forward spinoculation. After 3 days, the cells are infected with PRCV or SARS-CoV-2 EDB2 at an MOI of 0.01 and the supernatant is harvested at 24 HPI. Harvested samples are lysed in VL buffer and subjected to RT-qPCR for the determination of relative virus levels.

4.2.2 Identification of potential pan- and cross-species coronavirus restriction factors

To identify novel porcine ISGs that can restrict PRCV replication, pAPN-NSK cells were transduced with the porcine ISG library comprising 432 lentiviruses along with a negative control GFP lentivirus.

The screening experiment was performed in duplicate, using two aliquots of the library from the same production batch. However, based on RFP signal intensity, the transduction efficiency of screen 1 was higher than screen 2. Nonetheless, a relatively strong correlation was still observed, suggesting the overall quality of the aliquots used in both screens was preserved (**Fig4.6A**). It was not clear what caused the difference but screen 1 was performed shortly after the library was produced, where it was frozen at -80°C for less than a week. In contrast, screen 2 was performed more than a month later. We speculate the aliquot used for the second screen may have lost titre after being stored at -80°C for several weeks.

The screens were performed in eight 96-well plates with untransduced controls included in each plate for potential variability in infection levels across the plates. All samples were frozen once before subjecting to RT-qPCR to ensure consistency. Samples from plates 1 - 4 were processed one day post-freezing and samples from plates 5 - 8 were processed on day 2.

As expected, the majority of ISGs had a relatively limited impact on PRCV replication, with some of the most potent ISGs restricting replication by up to 32-fold (i.e. <-5

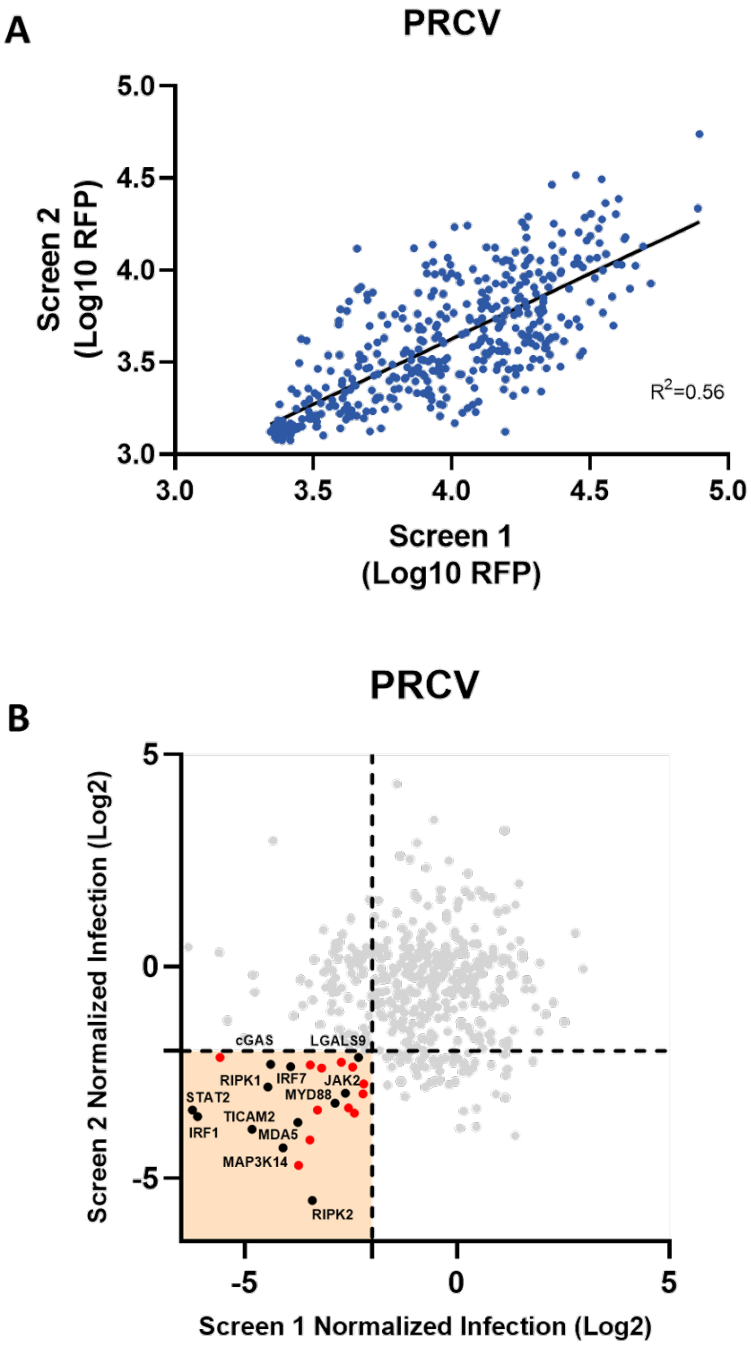


Figure 4.6: **Arrayed lentiviral porcine ISG screen against PRCV**. pAPN-NSK cells were seeded in 96-well plates and transduced with the porcine ISG library by forward spinoculation. **(A)** RFP fluorescence intensity was recorded daily and the readings on day 3 were reported. Transduced cells were infected with PRCV at an MOI of 0.01 on day 3 post-transduction with supernatants harvested 24 HPI and subjected to direct lysis RT-qPCR. **(B)** Infection levels of each ISG were normalised to the GFP negative control and reported as Log2 fold-change. The cut-offs were set at -2 (dotted lines) for both screens. Dots within the shaded area are considered hits. Dots in black represent ISGs that have a defined role in restricting virus replication; dots in red represent novel hits.

SSC	Gene	Summary
2	TBX3	Transcription factor involved in development
62	cGAS	PRR recognizing cytoplasmic DNA
73	TICAM2	Involved in NF-κB production
110	IRF1	Involved in IFN response
177	RIPK1	Involved in NF-κB production
186	STAT2	Involved in JAK/STAT signalling pathway
187	TREX1	3'-to-5' DNA exonuclease, known proviral factor
198	RAB8B	Small GTPases involved in vesicular trafficking
205	JAK2	Involved in JAK/STAT signalling pathway
212	ARHGAP42	Rho GTPase-activating protein involved in regulation of vascular tone
214	LACTB	Mitochondrial serine protease regulating mitochondrial lipid metabolism
215	NCOA1	Transcriptional coactivator for steroid and nuclear hormone receptors
218	ICAM1	Ligands for leukocyte adhesion
219	TEAD4	Transcription factor which plays a key role in the Hippo signalling pathway
220	ELF1	Transcription factor that activates the LYN and BLK promoters
233	MIS18BP1	Predicted to be involved in cell division
235	MYD88	Involved in NF-κB production
238	PIK3R3	Regulatory subunit of PI3K, coordinates a diverse range of cell functions
319	MAP3K14	Involved in NF-κB production
327	RIPK2	Involved in NF-κB production
403	PHACTR1	Phosphatase and actin regulator
409	LGALS9	Known Antiviral ISG, involved in apoptosis
416	MDA5	PRR recognizing dsRNA
429	IRF7	Involved in IFN response

Table 4.7: **Candidate PRCV restriction factors**. Complete list of hits identified in the PRCV screens (dots in the shaded area of Fig4.6). Genes shaded in blue represent ISGs that have a defined role in restricting virus replication. The function of each gene is summarised from Uniprot and GeneCard. SSC= Numerical identifier of genes in the library.

log2). A cut-off of -2 Log2 fold-change (i.e. 75% reduction) from the negative control GFP was set with ISGs below the cut-off in both replicates defined as hits. Among the 24 hits identified, half of them have defined roles in antiviral responses (**Fig4.6B and Table4.7**). Examples include the DNA and RNA sensors cGAS and MDA5; components of the NF-κB signalling pathway RIPK1, RIPK2, TICAM2, MYD88 and MAP3K14; components of the IFN and JAK/STAT signalling pathways IRF1, IRF7, STAT2 and JAK2; and a known antiviral ISG involved in apoptosis LGALS9. Some of the more novel hits such as TBX3 and ELF1 have also been previously described in other screens and suggested to have a potential board-spectrum or specific antiviral function [221, 292]. Re-identifying these hits has enhanced our confidence in the reliability of the current screens.

To test our hypothesis that viruses not adapted to replicating in pigs would be susceptible to additional ISGs, duplicate screens were performed using SARS-CoV-2 on hACE-2 NSK cells. The lentivirus library from the same production batch was used for the SARS-CoV-2 screens. The transduction efficiency between the two replicate SARS-CoV-2 screens was more consistent, with a higher correlation than the RFP signal for the PRCV screens (**Fig4.8A**). The RFP fluorescence intensity was also more comparable to the second PRCV screen and lower than the first, suggesting the relatively immediate application of the library after production does have a positive effect on transduction efficiency.

A total of 31 hits were identified in the SARS-CoV-2 screen, 17 of which have known characterised functions in reducing virus replication (**Fig4.8B and Table4.9**). Apart from the well-defined ISGs described above in the PRCV screen, additional well-characterised

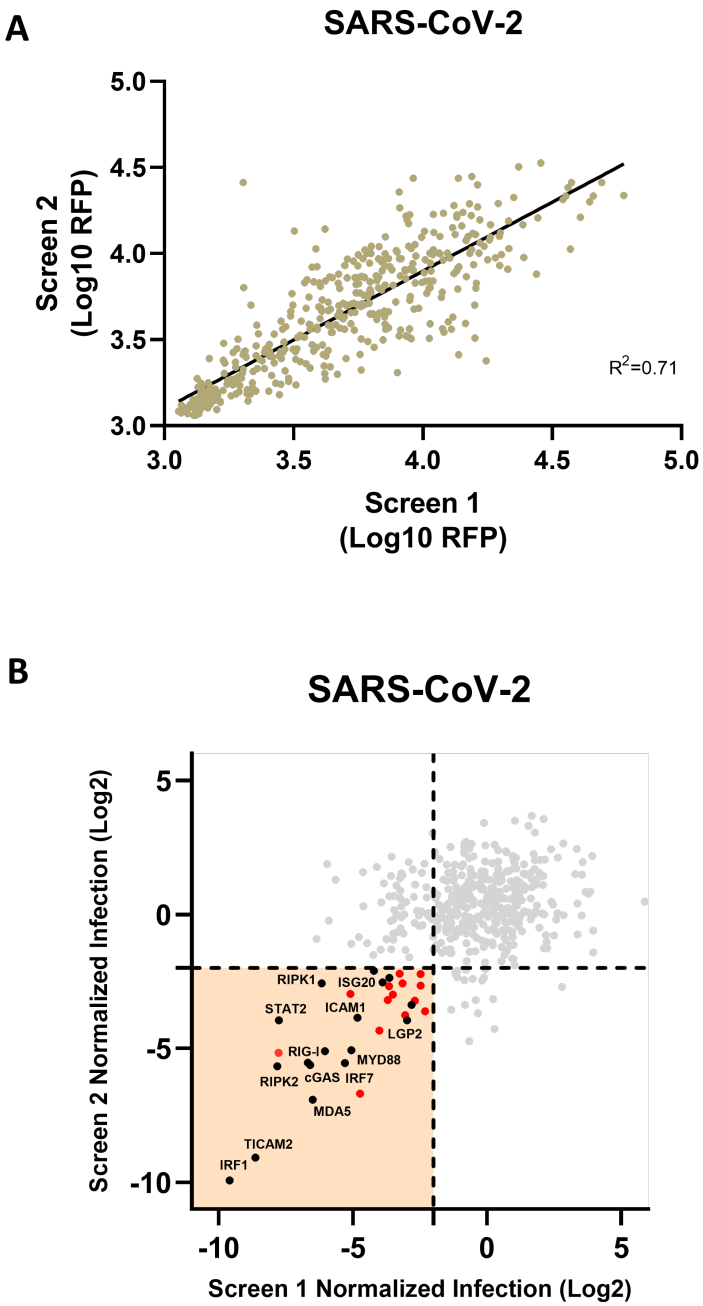


Figure 4.8: **Arrayed lentiviral porcine ISG screen against SARS-CoV-2**. hACE2-NSK cells were seeded in 96-well plates and transduced with the porcine ISG library by forward spinoculation. **(A)** RFP fluorescence intensity was recorded daily and the readings on day 3 were reported. Transduced cells were infected with SARS-CoV-2 isolate EDB2 at an MOI of 0.01 on day 3 post-transduction with supernatants harvested 24 HPI and subjected to direct lysis RT-qPCR. **(B)** Infection levels of each ISG were normalised to the GFP negative control and reported as Log2 fold-change. The cut-offs were set at -2 (dotted lines) for both screens. Dots within the shaded area are considered hits. Dots in black represent ISGs that have a defined role in restricting virus replication; dots in red represent novel hits.

SSC	Gene	Summary
2	TBX3	Transcription factor involved in development
8	STX11	Regulate protein transport between late endosomes and the trans-Golgi network
11	ISG20	Known Antiviral ISG
41	PLAG1	Transcription factor involved in proliferation, proto-oncogene
62	cGAS	PRR recognizing cytoplasmic DNA
73	TICAM2	Involved in NF-κB production
77	PML	Known antiviral factor by forming PML-nuclear body
86	RIG-I	PRR recognizing dsRNA
110	IRF1	Involved in IFN response
122	TRAF3IP2	Involved in NF-κB production
132	NRL	Transcription factor involved in the regulation of rod-specific genes
177	RIPK1	Involved in NF-κB production
186	STAT2	Involved in JAK/STAT signalling pathway
188	ZNF791-LIKE	Unknown
198	RAB8B	Small GTPases involved in vesicular trafficking
218	ICAM1	Ligands for leukocyte adhesion
235	MYD88	Involved in NF-κB production
244	NKX2-2	Transcription factor involved in the development of insulin-producing cells
256	PHACTR4	Phosphatase and actin regulator, regulates protein phosphatase 1
278	USF1	Transcription factor binding to E-box promoters
298	CCL11	Antimicrobial chemokine, promotes accumulation of eosinophils
327	RIPK2	Involved in NF-κB production
332	CD40	Receptor of tissue necrosis factor, involved in immunoglobin secretion
338	TRIM26	E3 ubiquitin-protein ligase, known to be anti or proviral depending on virus
361	CASP10	Involved in apoptosis
370	HOXD13	Transcription factor involved in morphogenesis
372	LGP2	PRR recognizing dsRNA
403	PHACTR1	Phosphatase and actin regulator
409	LGALS9	Known Antiviral ISG, involved in apoptosis
416	MDA5	PRR recognizing dsRNA
429	IRF7	Involved in IFN response

Table 4.9: **Candidate SARS-CoV-2 restriction factors**. Complete list of hits identified in the SARS-CoV-2 screens (dots in the shaded area of Fig4.8). Genes shaded in blue represent ISGs that have a defined role in restricting virus replication. The function of each gene is summarised from Uniprot and GeneCard. SSC= Numerical identifier of genes in the library.

hits, including two RNA sensors, RIG-I and LGP2, another factor involved in the NF-κB pathway TRAF3IP2, and the very well-studied antiviral ISG20 were also identified.

Some of these hits are also identified in published screening results against SARS-CoV-2 using the human ISG library, examples include RIPK1 and 2, STAT2, MYD88 and ISG20 [105, 107]. Some well-established restriction factors against SARS-CoV-2 such as LY6E and OAS1 were not identified in our screens as they are not included in the porcine ISG library. Similarly, some hits in our screens such as USF1, RAB8B, PHACTR1 and PHACTR4 are not included in the human library as they are not considered ISGs in humans [154].

To select hits for further characterisation, the results from the PRCV and SARS-CoV-2 duplicate screens were averaged and plotted against each other. The same cut-off was used and genes that restrict both viruses by more than 75% were considered pan-coronavirus restriction factor candidates (**Fig4.10A and Table4.11**). Among the 15 better-characterised hits (i.e. black dots), IRF1 is the most potent ISG for both viruses, reducing SARS-CoV-2 and PRCV replication by nearly 1000-fold and 30-fold respectively. The difference in maximum restriction potency (-10 vs -5 Log2 fold-change) is consistent with our hypothesis and previous experiments, with SARS-CoV-2 being more sensitive to porcine IFN in NSK-derived cells than PRCV. Generally speaking, ISGs involved in IFN signalling (i.e. IRF1, STAT2, cGAS, MDA5 etc.) and the NF-κB pathway (i.e. TICAM2, RIPK1, RIPK2, MYD88 etc.) are among the most potent hits in both screens. This result is consistent with previous human ISG screens where these ISGs were found to trigger IFN production through overexpression alone and restrict a broad range of pathogens.

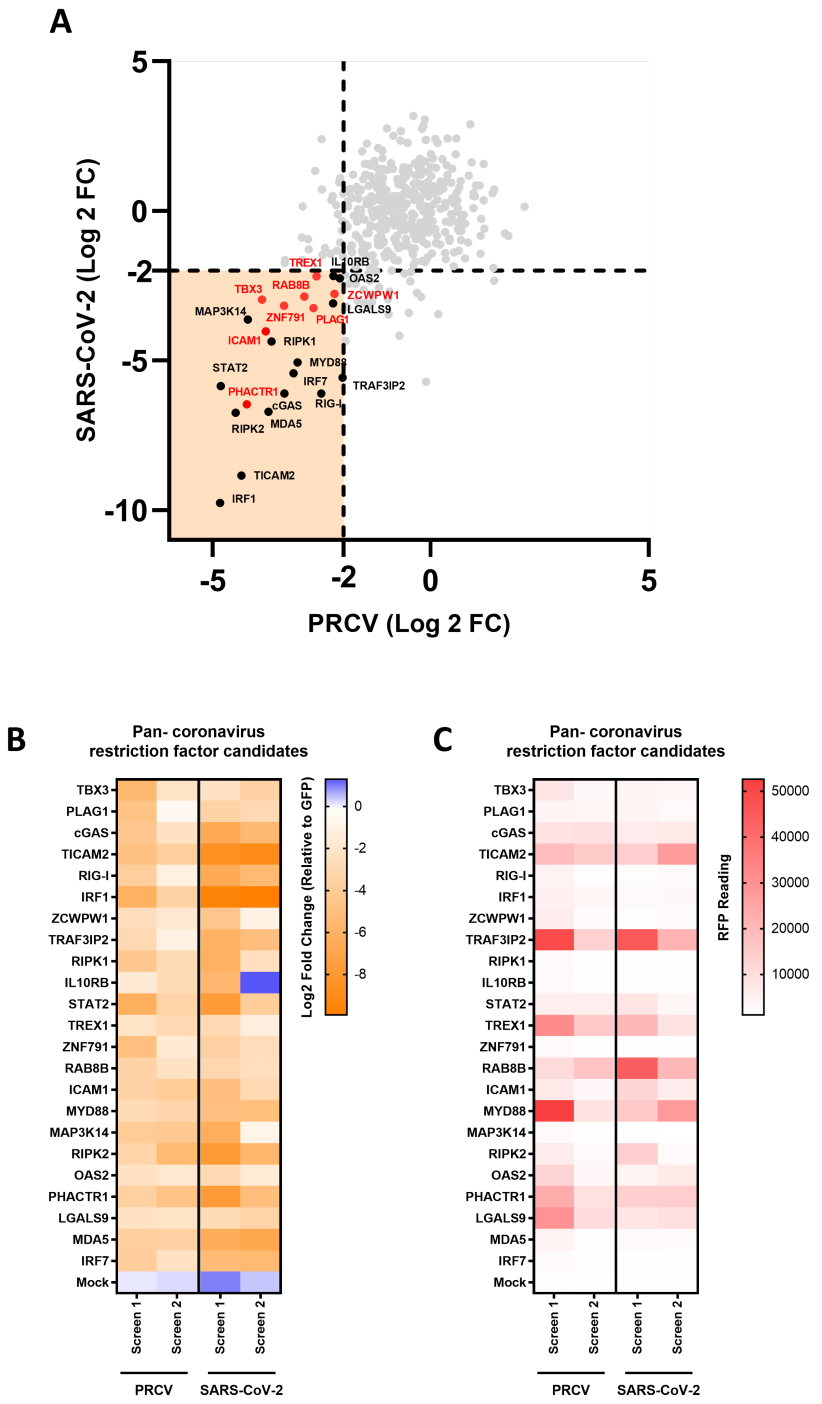


Figure 4.10: **Comparative analysis of PRCV and SARS-CoV-2 screens reveals potential pan-coronavirus restriction factors**. **(A)** The biological duplicates of the PRCV and SARS-CoV-2 screens **Fig4.6B and Fig4.8B** were averaged and plotted against each other. The cut-offs were set at -2 (dotted lines) for both viruses. Dots within the shaded area are considered hits. Dots in black represent ISGs that have a defined role in restricting virus replication; dots in red represent novel hits. **(B)** The Log2 fold-change and **(C)** RFP fluorescence intensity of each candidate ISG in each screen were shown by a heat map.

SSC	Gene	Summary
2	TBX3	Transcription factor involved in development
41	PLAG1	Transcription factor involved in proliferation, proto-oncogene
62	cGAS	PRR recognizing cytoplasmic DNA
73	TICAM2	Involved in NF-κB production
86	RIG-I	PRR recognizing dsRNA
110	IRF1	Involved in IFN response
113	ZCWPW1	Dual histone methylation reader essential for spermatogenesis
122	TRAF3IP2	Involved in NF-κB production
177	RIPK1	Involved in NF-κB production
181	IL10R2	Involved in JAK/STAT signalling pathway
186	STAT2	Involved in JAK/STAT signalling pathway
187	TREX1	3'-to-5' DNA exonuclease, proviral factor for HIV and IAV
188	ZNF791-like	Unknown
198	RAB8B	Small GTPases involved in vesicular trafficking
218	ICAM1	Ligands for leukocyte adhesion
235	MYD88	Involved in NF-κB production
319	MAP3K14	Involved in NF-κB production
327	RIPK2	Involved in NF-κB production
386	OAS2	Known Antiviral ISG
403	PHACTR1	Phosphatase and actin regulator
409	LGALS9	Known Antiviral ISG, involved in apoptosis
416	MDA5	PRR recognizing dsRNA
429	IRF7	Involved in IFN response

Table 4.11: **Candidate pan-coronavirus restriction factors**. Complete list of potential pan-coronavirus restriction factors (dots in the shaded area of Fig4.10). Genes shaded in blue represent ISGs that have a defined role in restricting virus replication. The function of each gene is summarised from Uniprot and GeneCard. SSC= Numerical identifier of genes in the library.

ISGs with direct antiviral effects such as OAS2 and LGALS9 are also identified. Among the 8 novel hits (i.e. red dots), PHACTR1 exhibits the greatest potency against both viruses. The less-characterised hits were selected for further validation.

To assess the effect of transduction efficiency on the relative potency of individual hits against PRCV and SARS-CoV-2, the Log2 fold-change and RFP fluorescence intensity of each pan-coronavirus restriction factor candidate for each screen are shown (**Fig4.10B and C**). However, the majority of hits have a relatively low fluorescence intensity, consistent with results shown earlier where RFP intensity does not necessarily correlate with GFP (i.e. ISG) expression.

Apart from identifying novel porcine pan-coronavirus restriction factors, one of the major aims of this project is to identify restriction factors that can restrict SARS-CoV-2 more potently than PRCV, indicating a potential role in species-specific inhibition of virus replication. An arbitrary cut-off of $\pm$ 1 Log2 fold-change in PRCV and -2 Log2 fold change in SARS-CoV-2 were set and ISGs falling under these criteria were defined as cross-species coronavirus restriction factors (**Fig4.12A and Table4.13**).

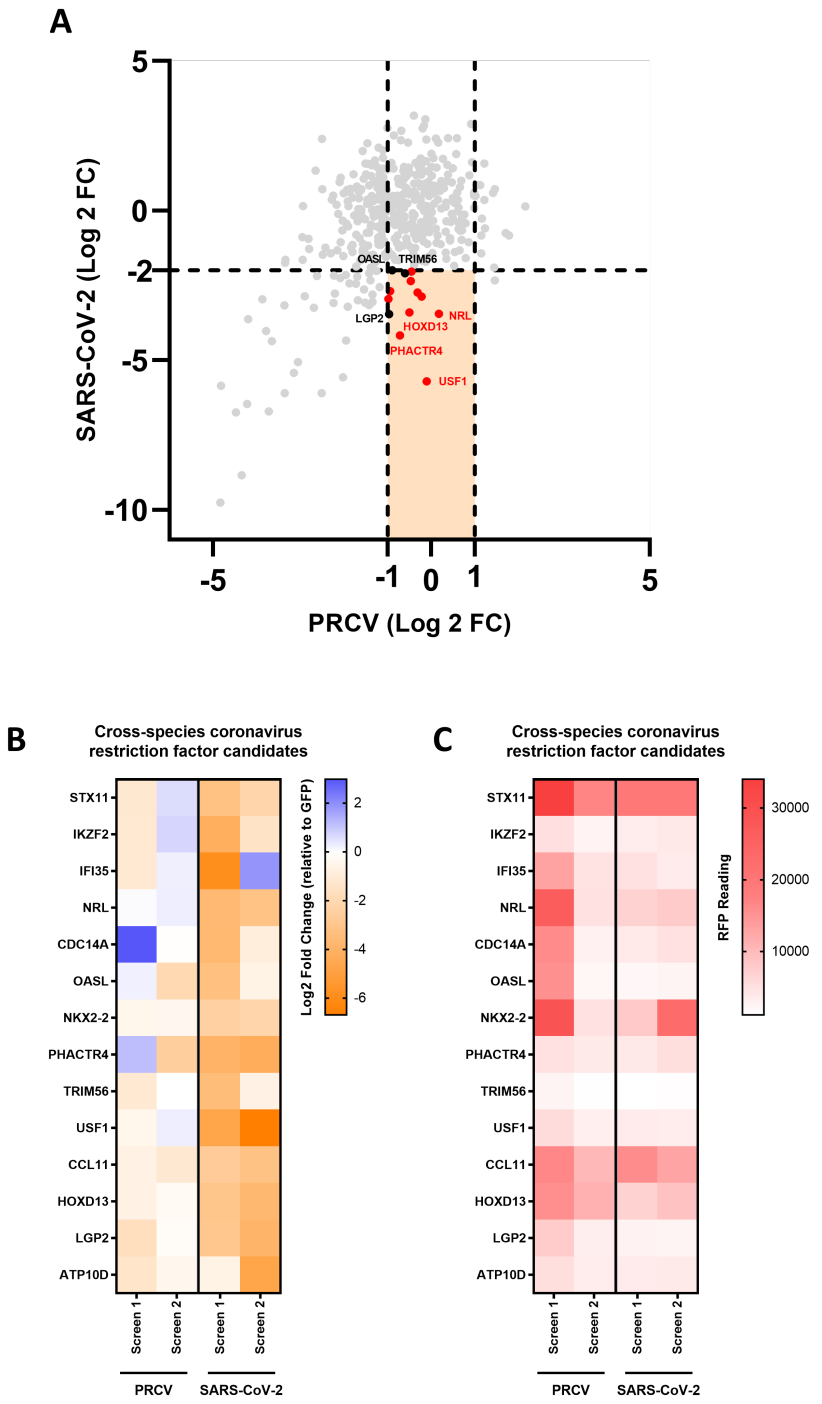


Figure 4.12: **Comparative analysis of PRCV and SARS-CoV-2 screens reveals potential cross-species coronavirus restriction factors.** **(A)** The biological duplicates of the PRCV and SARS-CoV-2 screens **Fig4.6B and Fig4.8B** were averaged and plotted against each other. The cut-offs were set at ± 1 for PRCV and -2 for SARS-CoV-2 (dotted lines). Dots within the shaded area are considered hits. Dots in black represent ISGs that have a defined role in restricting virus replication; dots in red represent novel hits. **(B)** The Log2 fold-change and **(C)** RFP fluorescence intensity of each candidate ISG in each screen were shown by a heat map.

SSC	Gene	Summary
8	STX11	Regulate protein transport between late endosomes and the trans-Golgi network
39	IKZF2	Transcription factors involved in the regulation of lymphocyte development
105	IFI35	Regulate production of Type I IFN, known to be anti- or proviral depending on virus
132	NRL	Transcription factor involved in the regulation of rod-specific genes
147	CDC14A	Regulate the function of p53, involve in mitosis
178	OASL	Known Antiviral ISG
244	NKX2-2	Transcription factor involved in the development of insulin-producing cells
256	PHACTR4	Phosphatase and actin regulator, regulate protein phosphatase 1
275	TRIM56	Involved in cGAS-STING and other innate immune pathways
278	USF1	Transcription factor binding to E-box promoters
298	CCL11	Antimicrobial chemokine, promotes accumulation of eosinophils
370	HOXD13	Transcription factor involved in morphogenesis
372	LGP2	PRR recognizing dsRNA
428	ATP10D	Predicted to be involved in phospholipid translocation

Table 4.13: **Candidate cross-species coronavirus restriction factors**. Complete list of potential species-specific coronavirus restriction factors (dots in the shaded area of Fig4.12). Genes shaded in blue represent ISGs that have a more defined role in restricting virus replication. The function of each gene is summarised from Uniprot and GeneCard. SSC= Numerical identifier of genes in the library.

Similarly, the Log2 fold-change and RFP fluorescence intensity of each cross-species coronavirus restriction factor candidate for each screen are shown (**Fig4.12B and C**). The overall RFP intensities of these hits are higher compared to the pan-coronavirus restriction factors, suggesting their antiviral potency might be more dose-dependent. Among the 14 hits, 3 have a defined antiviral role. LGP2, OASL and TRIM56 are all less efficient in restricting PRCV than SARS-CoV-2, although the difference is not large. Other hits such as USF1, NRL, PHACTR4 and HOXD13 displayed a larger differential restriction inhibition, all inhibiting SARS-CoV-2 to a greater extent than PRCV.

4.2.3 Preliminary validation of selected candidate ISGs

To validate the pan- and cross-species candidate ISGs, high-titre lentiviruses were produced in T75 culture flasks. Initial attempts to achieve similar levels of transduction were unsuccessful as the variation in titres between ISG lentivirus was too high, possibly due to the impact of the ISGs on lentiviral production itself, and in some cases impacts on RFP signal. Instead, we adjusted the lentivirus input to generate high levels of transduction, while remaining below 50,000 fluorescence units, the threshold at which non-specific inhibition was observed in previous experiments (**Fig4.3G and H**). This maximised the potential for detecting the phenotype of individual ISGs while ensuring the effect is not due to excessive transduction.

To test if the results from the screening experiments are reproducible. pAPN- or hACE2-NSK cells were seeded in 96-well plates and transduced with high-titre lentivirus expressing indicated ISGs. All procedures and conditions were identical to the screens. The results showed that most pan-coronavirus restriction factor candidates reduced

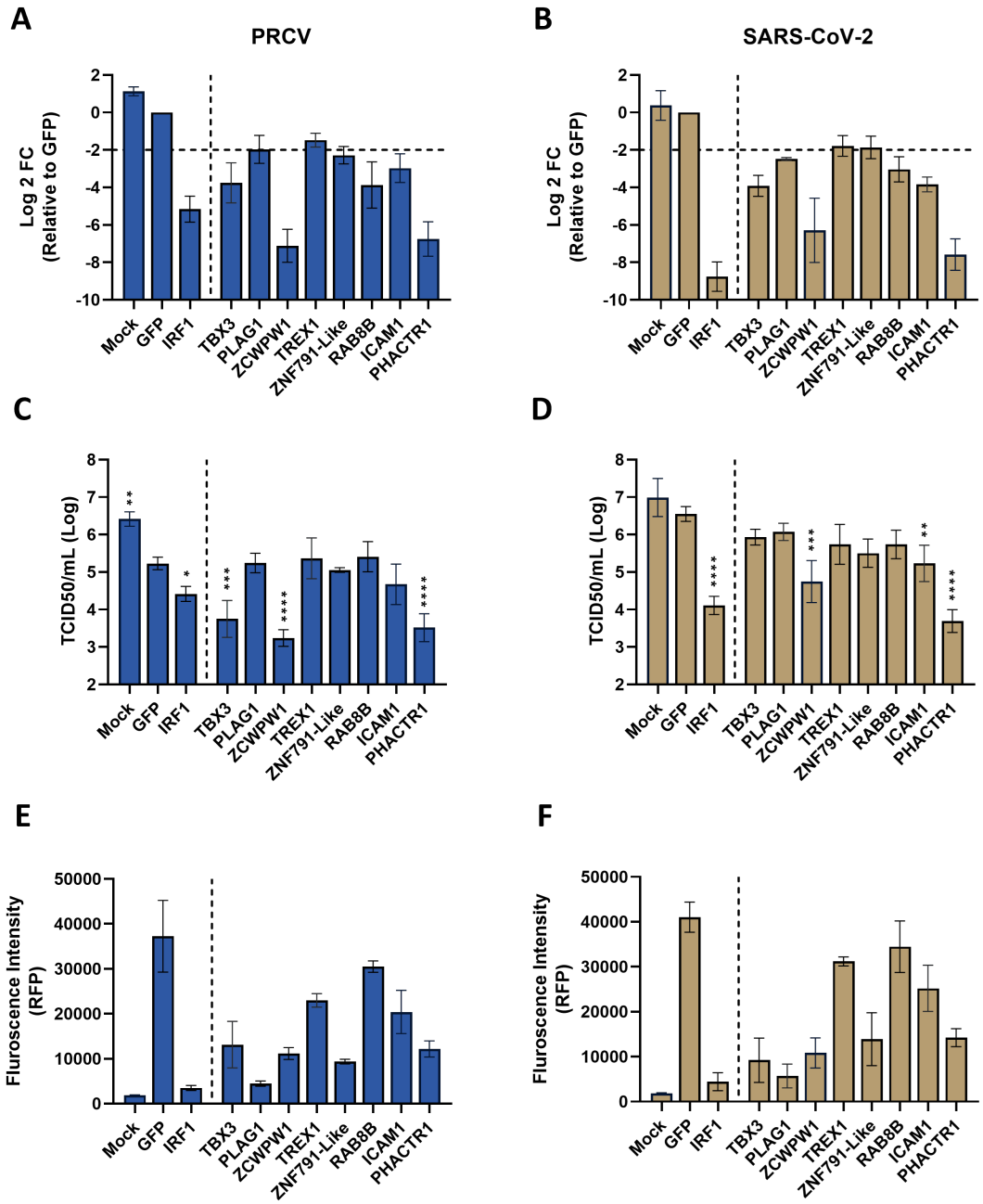


Figure 4.14: **Validation of pan-coronavirus restriction factors candidates**. pAPN- or hACE2-NSK cells were seeded in 96-well plates and transduced with high-titre lentivirus expressing indicated ISGs generated from T75 culturing flasks. All other procedures and conditions are identical to previous screening experiments. Supernatant samples were harvested for direct lysis RT-qPCR for **(A)** PRCV and **(B)** SARS-CoV-2. The same samples were used for **(C)** PRCV titring on ST cells or **(D)** SARS-CoV-2 titring on hACE2-NSK. The transduction efficiencies of indicated ISGs in **(E)** pAPN-NSK and **(F)** hACE2-NSK were recorded daily and readings on day 3 were reported. Results represent the average of three independent experiments, each with two technical replicates. Error bars represent the standard deviation. Multiple comparisons were performed from a one-way ANOVA against the control GFP.****= $p<0.0001$, ***= $p<0.001$ **= $p<0.01$, *= $p<0.05$.

PRCV and SARS-CoV-2 replication by at least 75% by direct lysis RT-qPCR (**Fig4.14A and B**), thereby validating the original screens. While RT-qPCR was an effective method for determining virus replication for high throughput screens, it is not a direct measure of virus production. Therefore, we sought to directly test infectious virus production using TCID50 assays. Surprisingly, the reduction in virus levels was consistently higher when measured by TCID50 compared to RTqPCR (**Fig4.14C**). The negative control GFP had reduced PRCV titre by more than 10-fold when compared to untransduced cells. This diminished the measured effect of other ISG as demonstrated by the potent ISG IRF1, reducing PRCV titre by only half a log. Nonetheless, with enhanced transduction efficiency compared to the screens, TBX3, ZCWPW1 and PHACTR1 significantly reduced PRCV titre by approximately one to two logs.

In contrast to PRCV, reduced viral replication due to transduction alone was not observed with SARS-CoV-2 infections in hACE2-APN cells, consistent with RT-qPCR results (**Fig4.14C**). Although not statistically significant, all candidate ISGs reduced SARS-CoV-2 titre to varying extents. Besides IRF1 reducing virus titre by more than 100-fold, ZCWPW1, PHACTR1 and ICAM1 also significantly reduced virus titres, with PHACTR1 being even more potent than IRF1. The transduction efficiencies of all ISGs were below GFP, confirming the reduction in virus titre is not an unspecific response to lentiviral transduction (**Fig4.14E and F**).

Similarly, we have picked three candidates from the list of potential cross-species coronavirus restriction factors for initial validation. USF1, NRL and HOXD13 were selected as they showed consistent results across all 4 screens, reducing SARS-CoV-2 replication

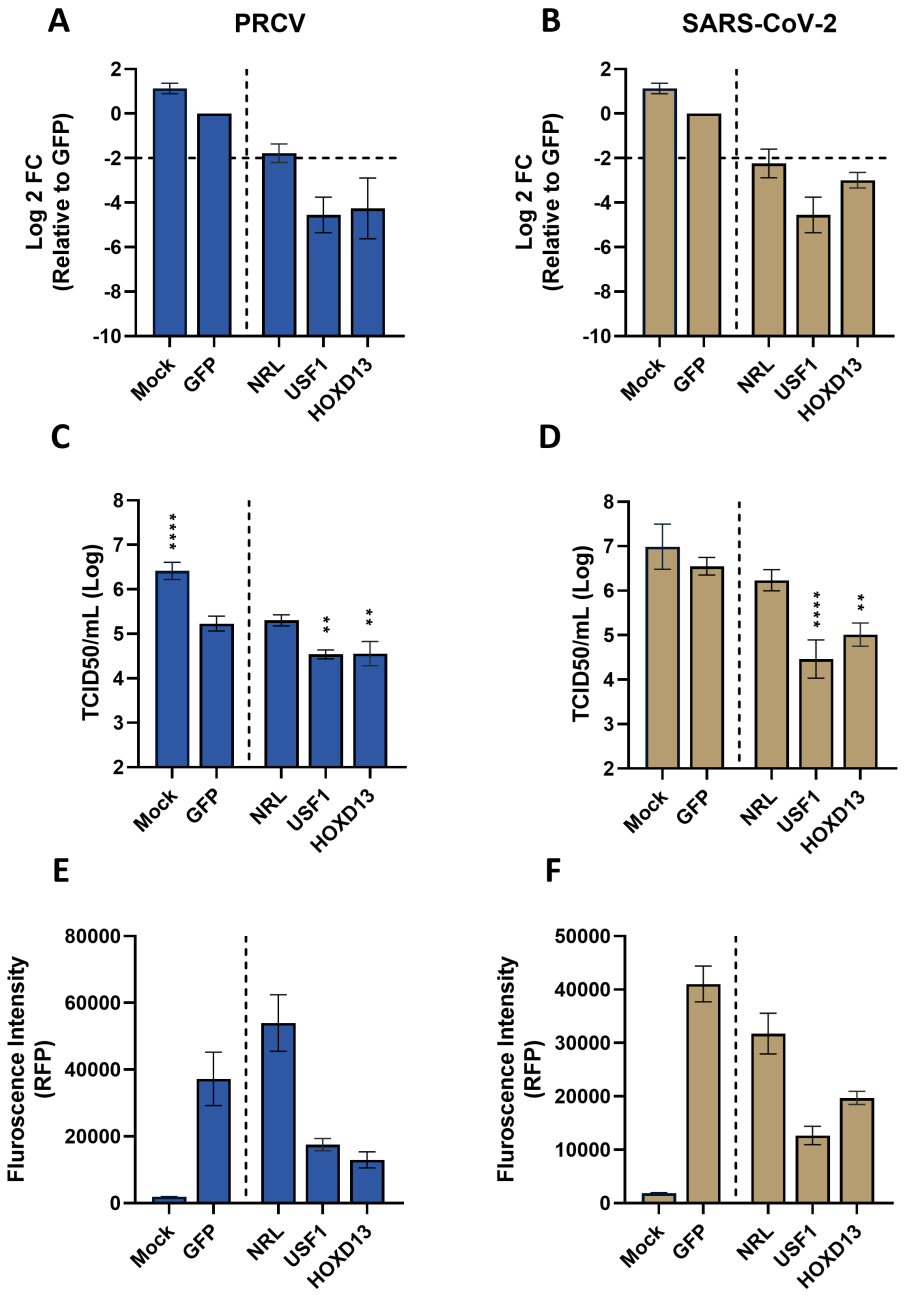


Figure 4.15: **Validation of cross-species coronavirus restriction factors candidates**. pAPN- or hACE2-NSK cells were seeded in 96-well plates and transduced with high-titre lentivirus expressing indicated ISGs generated from T75 culturing flasks. All other procedures and conditions were identical to previous screening experiments. Supernatant samples were harvested for direct lysis RT-qPCR for **(A)** PRCV and **(B)** SARS-CoV-2. The same samples were used for **(C)** PRCV titring on ST cells or **(D)** SARS-CoV-2 titring on hACE2-NSK. The transduction efficiencies of indicated ISGs in **(E)** pAPN-NSK and **(F)** hACE2-NSK were recorded daily and readings on day 3 were reported. Results represent the average of three independent experiments, each with two technical replicates. Error bars represent the standard deviation. Multiple comparisons were performed from a one-way ANOVA against the control GFP.****= $p<0.0001$, **= $p<0.01$.

while having a modest effect on PRCV. Consistent with the screening result, these hits were able to reduce SARS-CoV-2 replication by at least 75% when measured by direct lysis RT-qPCR. However, unlike the screens, they also show potent restriction of PRCV replication (**Fig4.15A and B**). As mentioned above, despite not being observed by RT-qPCR, GFP transduction alone significantly reduced PRCV titres by more than 10-fold, causing USF1 and HOXD13 to have reduced infectious virus titres by approximately half a log. Nonetheless, SARS-CoV-2 titres were significantly reduced by 50 to 100-fold (**Fig4.15C and D**). Despite having a high transduction efficiency (higher than GFP in the case of PRCV), NRL does not seem to affect the infectious virus titre of either virus, indicating a failure to validate the hit in this case (**Fig4.15E and F**).

The validation results of the cross-species candidates lead us to speculate that rather than being able to completely evade, PRCV merely has a higher tolerance than SARS-CoV-2 against these ISGs. As the lentiviruses used in the validation experiments had a higher titre, and therefore higher transduction efficiency, than in the screening experiment it is possible that PRCV can only counteract the expression level of said ISGs in the screen but not in the following experiments. To address this, cells were transduced with 2-fold serially diluted USF1 lentivirus to generate cells with different expression levels. Nonetheless, we failed to replicate the results we observed from the screening experiment, as reducing the lentivirus input diminished the effect of USF1 on both viruses (**Fig4.16A and B**). However, USF1 consistently shows a greater inhibition of SARS-CoV-2 replication than PRCV.

Thus far, all infection experiments were performed in cell lines overexpressing their

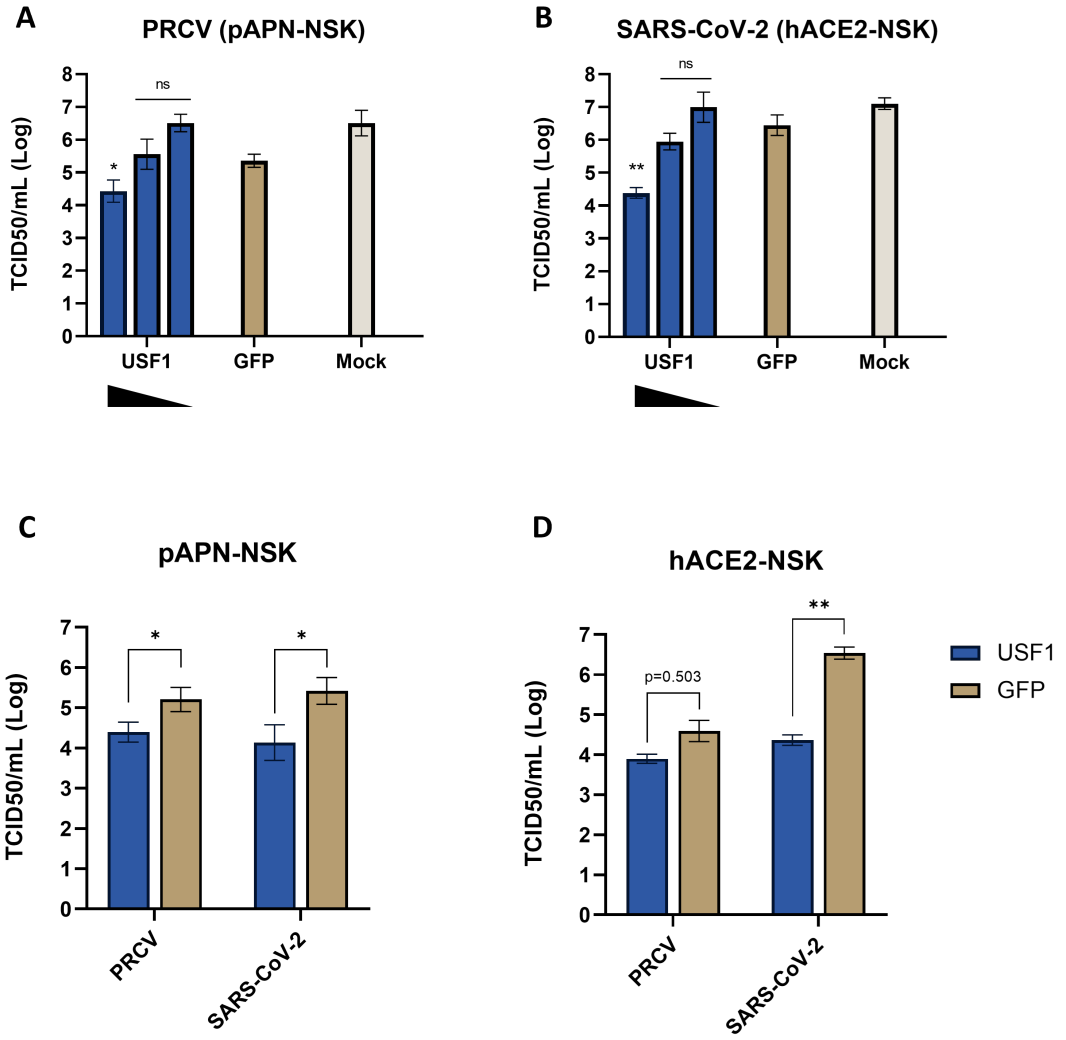


Figure 4.16: **Differential effect of USF1 correlates with virus replication kinetics.** Cells were seeded in 96-well plates and transduced with 2-fold serially diluted USF1 lentivirus or GFP. **(A)** pAPN- and **(B)** hACE2-NSK cells were infected with PRCV or SARS-CoV-2 at an MOI of 0.01 respectively. Similarly, **(C)** pAPN- or **(D)** hACE2-NSK cells were transduced with USF1 or GFP. Transduced cells were infected with either PRCV or SARS-CoV-2 at an MOI of 0.01. Supernatant samples were harvested for titring, PRCV on ST cells and SARS-CoV-2 on hACE2-NSK cells. Results represent the average of two independent experiments, each with two technical replicates. Error bars represent the standard deviation. Unpaired t-tests were used to calculate statistical significance between USF1 and GFP. **= $p<0.01$, *= $p<0.05$

respective entry receptors (i.e. SARS-CoV-2 on hACE2-NSK and PRCV on pAPN-NSK). The observed phenotype of USF1 might be an artefact of SARS-CoV-2 being able to replicate to a higher titre than PRCV (10^7 vs 10^6) and thus a wider dynamic range for reduction. To address this, the experiment was repeated in both cell lines with both viruses. pAPN- and hACE2- NSK cells were seeded in 96-well plates and transduced with either USF1 or GFP lentivirus. Transduced cells were then infected by either PRCV or SARS-CoV-2. Indeed, the differential potency of USF1 towards the two viruses seems to correlate with the virus titre. In pAPN-NSK, where the titre of SARS-CoV-2 is similar to PRCV, the restriction potency of USF1 on both viruses is similar. In contrast, in hACE2-NSK, where the titre of SARS-CoV-2 is substantially higher than PRCV, USF1 restricts SARS-CoV-2 by more than 100-fold when compared to less than 10-fold with PRCV (**Fig4.16C and D**).

The above data suggest the cross-species coronavirus restriction factors identified in the screens are possibly an artefact of differential virus replication dynamics. We have therefore shifted our focus to validating the pan-coronavirus restriction factors. PHACTR1 showed the most potent effect in reducing coronavirus replication from the validation experiments (**Fig4.14C and D**). Intriguingly, this gene was also identified as a hit in separate screening experiments against IAV and has been validated to restrict a broad range of viruses. This gene, along with the closely related PHACTR4 (a cross-species coronavirus restriction factor candidate), is currently being characterised by another member of our group. The second most potent pan-coronavirus restriction factor candidate ZCWPW1 has therefore been selected for further characterisation.

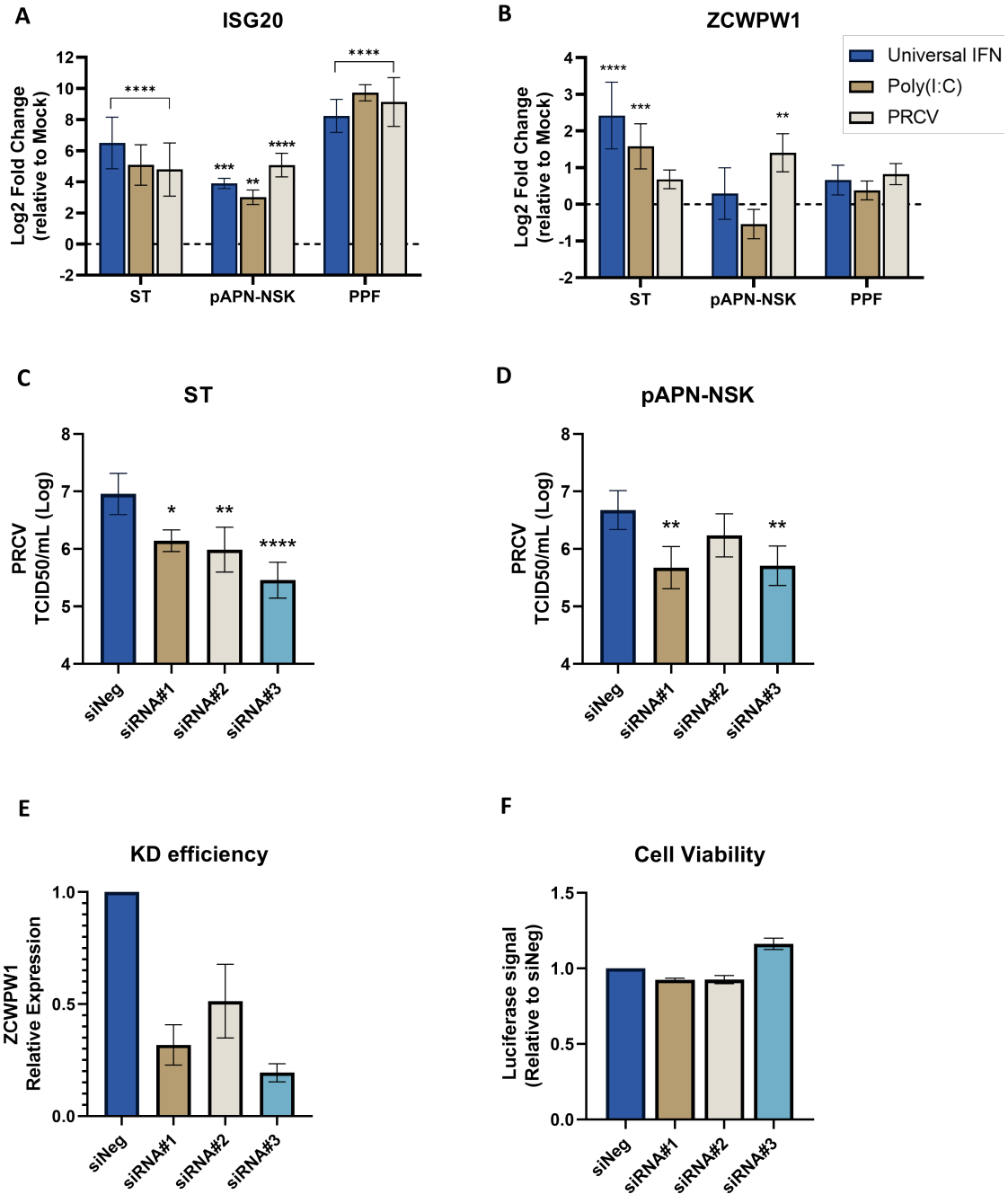


Figure 4.17: **Initial characterisation of a potential pan-coronavirus restriction factor ZCWPW1.**

Figure 4.17: **Initial characterisation of a potential pan-coronavirus restriction factor ZCWPW1.** Indicated cells were seeded in 24-well plates and treated with 1000U/ml universal IFN, transfected with 5ug/ml poly(I:C) or infected with PRCV at an MOI of 0.01. Cell lysate samples were harvested 24 hours later and RNA was extracted for the measurement of **(A)** ISG20 and **(B)** ZCWPW1 expression by RT-qPCR. **(C)** ST or **(D)** pAPN-NSK cells were reverse-transfected in 24-well plates with siNeg or siRNA targeting ZCWPW1. Cells were PRCV-infected 4 days post-transfection and samples were harvested and titred 24 HPI. **(E)** Expression of ZCWPW1 in transfected ST cells was measured by RT-qPCR. **(F)** ST cells were seeded in 96-well plates and transfected. A luciferase-based assay was used to measure cell viability. All gene expressions were normalized to GAPDH. Results represent the average of three independent experiments, each with two technical replicates. Error bars represent the standard deviation. Multiple comparisons were performed from a one-way ANOVA. Comparisons were against Mock for **(A)** and **(B)** and siNeg for **(C)** and **(D)**. ****=$p<0.0001$, ***=$p<0.001$, **= $p<0.01$, *= $p<0.05$. siNeg= Negative control siRNA, KD=knock down.

As our ISG library was generated based on published RNA-sequencing data on IFN-treated porcine primary cells, we first wanted to test if ZCWPW1 is IFN-inducible. ST, pAPN-NSK, and porcine primary fibroblast (PPF) cells generated in-house were treated with universal IFN, transfected with polyinosinic: polycytidylic acid (Poly(I:C)) or infected by PRCV. Cell lysate samples were harvested for RNA extraction 24 hours later and the expressions of ZCWPW1, as well as the known IFN-inducible ISG20, were measured by RT-qCR (**Fig4.17A and B**). As expected, ISG20 was upregulated significantly in all 3 cell types under all 3 conditions. The largest response was observed in PPF cells, with increased ISG20 expression close to 1000-fold under all conditions. In contrast, a relatively weak response towards the stimuli was observed in pAPN-NSK cells, where expression increased by approximately 10 to 50-fold (**Fig4.17A**). While the induction of ZCWPW1 was lower than ISG20, a significant increase in expression was observed in ST cells treated with universal IFN or transfected with Poly(I:C). A subtle increase was also observed in PRCV-infected pAPN-NSK cells. However, no induction was observed in PPF cells in any of the conditions (**Fig4.17B**).

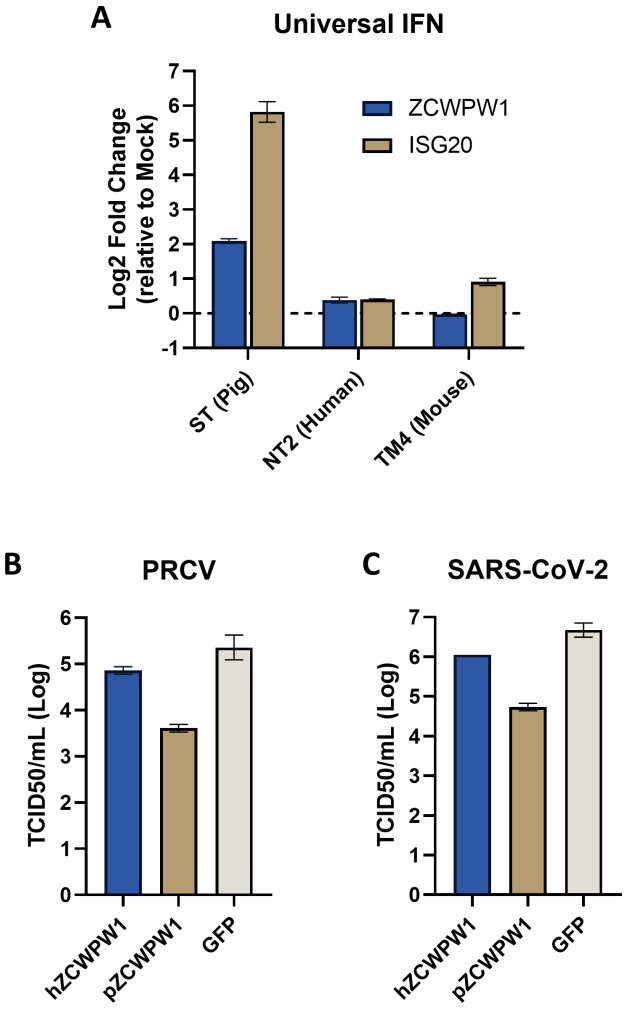


Figure 4.18: **Preliminary characterisation of ZCWPW1 from other species**. **(A)** Indicated cells were seeded in 24-well plates and treated with 1000U/ml universal IFN. Cell lysate samples were harvested 24 hours later and RNA was extracted for the measurement of ISG20 and ZCWPW1 expression by RT-qPCR. All gene expressions were normalized to GAPDH. n=2x2, error bars represent the standard deviation from 2 independent experiments. **(B)** pAPN- and **(C)** hACE2-NSK cells were seeded in 96-well plates and transduced with human or pig ZCWPW1 or GFP. Transduced cells were infected by PRCV and SARS-CoV-2 at an MOI of 0.01 respectively. Supernatant samples were harvested for titring, PRCV on ST and SARS-CoV-2 on hACE2-NSK. Results represent the average of two technical replicates. Error bars represent the standard deviation. hZCWPW1= human ZCWPW1; pZCWPW1= porcine ZCWPW1.

To further characterise ZCWPW1, 3 siRNAs targeting different regions of the coding sequence of porcine ZCWPW1 mRNA were designed and prepared. As overexpression of this gene exhibits an antiviral phenotype in the target cell, it was hypothesized that siRNA knockdown would result in increased virus replication. The designed siRNAs were transfected into ST and pAPN-NSK cells (as they both show a trend of increased expression in response to stimuli) and challenged with PRCV. Surprisingly, instead of showing an opposite phenotype, the knockdown of ZCWPW1 produced a similar phenotype as overexpression, reducing infectious virus titre significantly by 10 to 50-fold (**Fig4.17C and D**). The level of restriction observed in the 3 siRNAs also partially correlates with their respective knockdown efficiency, siRNA#3 has the strongest efficiency, showing the strongest phenotype in ST cells and an equally strong phenotype as siRNA#1 in pAPN-NSK cells. In contrast, siRNA#2 has the poorest efficiency, showing a weaker phenotype in ST cells (**Fig4.17E**). Knockdown of ZCWPW1 did not impact cell viability, suggesting the reduction in virus titre is not due to toxicity associated with the siRNAs (**Fig4.17F**).

To determine if the interferon-inducible feature of ZCWPW1 in testis cells is conserved in other species, NT2 (epithelial-like human testis carcinoma) and TM4 (mouse Sertoli cells) were treated with universal IFN and expression of ZCWPW1 and ISG20 were measured as mentioned above. ZCWPW1 did not upregulate in NT2 or TM4 cells following IFN treatment. However, surprisingly, the induction of marker ISG20 in both cells is very subtle, with less than a 2-fold increase compared to a nearly 50-fold increase in ST cells (**Fig4.18A**). This suggests the constant level of ZCWPW1 following treatment may be attributed to the unresponsive nature of the cells.

To determine if the antiviral function of ZCWPW1 is conserved in humans, the coding sequence of human ZCWPW1 (Ensembl ID: ENST00000684423) is acquired commercially (Invitrogen) and cloned into pSCRPSY. Lentivirus was produced and used for transduction of pAPN- or hACE2-NSK cells, followed by infection with PRCV and SARS-CoV-2 respectively. Despite achieving a similar transduction level as porcine ZCWPW1, human ZCWPW1 showed no significant effect on reducing PRCV or SARS-CoV-2 replication (**Fig4.18B and C**).

4.3 Discussion

Previous human ISG screens had identified a wide range of antiviral factors restricting a specific or broad range of viruses, deepening our understanding of virus-host interactions in humans. As the One Health concept suggests, animal health and human health are interdependent. We have therefore expanded the scope of these studies to pigs, developing a porcine ISG library aiming to understand virus-host interactions in this species. In this chapter, we have applied the library to screen against PRCV and SARS-CoV-2, identifying some novel porcine-specific restriction factors along with some previously characterised antiviral factors in humans, demonstrating both divergent and conserved IFN responses against coronaviruses between humans and pigs.

The screens identified many porcine ISGs that have been characterised to have antiviral functions against a broad range of viruses in humans (**Fig4.10B and Table4.11**). Examples include some PRRs (cGAS, RIG-I and MDA5) and components in the IFN signalling pathway (IRFs and STATs) and NF-κB pathway (RIPKs and MyD88). This is consistent with reports suggesting that, although the extracellular (ligand-bind) regions of PRRs are under strong positive selection, regions responsible for their downstream signalling pathway are generally conserved [293]. The data suggests that the overexpression of these porcine ISGs likely resulted in the expression of IFN or other pro-inflammatory cytokines, similar to what was observed in their human counterpart, and they are indeed ideal positive controls to be included in the ISG library. Nonetheless, the major aim for this project and overexpression screens is to identify direct antiviral ISGs, and we have identified and validated some potent porcine restriction factors against coronaviruses,

such as PHACTR1 and ZCWPW1 that have not been previously described in human screens.

The initial characterisation of ZCWPW1 has yielded some interesting results. This Zinc Finger CW-Type and PWWP Domain Containing 1 protein is poorly characterised. While predicted to enable methyl-CpG and methylated histone binding, recent studies in mouse models suggest it is involved in male fertility and spermatogenesis. According to the Genotype-Tissue Expression (GTEx) Portal (https://www.gtexportal.org) and several published reports, the expression of this gene and its protein product is highly enriched in testis [294–296]. Coincidentally, we have observed a robust induction of ZCWPW1 in ST cells, which is of testis origin. To investigate if this tissue-specific induction is conserved across species, the expression of ZCWPW1 in a human and mouse testis cell line following universal IFN treatment was measured. However, these cell lines do not seem to respond to IFN, as the induction of the marker ISG20 was very subtle (**Fig4.18A**). Testing on other IFN-responsive testis cell lines is needed before a conclusion can be made. Induction of ZCWPW1 was not observed in other IFN-responsive human cell lines including NHDF and A549 cells (data not shown). The amino acid sequence similarity between human, mouse and porcine ZCWPW1 is less than 70% and they could potentially have different molecular functions. Indeed, preliminary data suggests human ZCWPW1 has little to no antiviral activity in pAPN- or hACE2-NSK cells against PRCV or SARS-CoV-2 (**Fig4.18B and C**). However, due to time constraints, this experiment has only been performed once and more repeats are needed to confirm this species-specific difference. More experiments are also needed to determine whether this is due to the differences in amino acid sequence or incompatibility of human factors

in a porcine cell line. This includes expressing the porcine and human ZCWPW1 in either their respective or opposite cell background and testing for their antiviral potency. Generation of chimeric porcine and human ZCWPW1 can also help determine the specific region or sequence responsible for the potential porcine-specific antiviral potency.

To date, few studies have been dedicated to examining the ISG profile of pigs. According to Shaw *et al.*, ZCWPW1 is potentially a porcine-specific ISG, as induction following IFN response was not observed in other studied species [154]. ZCWPW1 has never been described as an ISG in any human studies or database, meanwhile, several porcine studies have recorded ZCWPW1 mRNA induction following IFN treatment or virus infection [297, 298]. In two of the studies, a relatively large induction was observed, ranging from 20 to 100-fold compared to 4-fold in our study. However, cells were harvested 4 or 8 hours after IFN treatment instead of 24 hours. We speculate induction of ZCWPW1 is more potent early on the IFN response. Repeating the experiment in **Fig4.17A and B** at different time points could capture a more complete picture of ZCWPW1 induction.

It was initially surprising that both siRNA knockdowns and overexpression of ZCWPW1 restrict virus replication. However, this has been reported in other well-characterised antiviral ISGs. For example, there is contradicting evidence on whether IFITMs restrict or support SARS-CoV-2 entry from overexpression and knockdown studies. It was later suggested that endogenous IFITM expression does support efficient infection of SARS-CoV-2, while artificial overexpression may have changed the topology, localization or endocytic activity of IFITMs, thereby exhibiting an unnatural restrictive phenotype [148]. It is possible that a similar situation exists with ZCWPW1. Another possible

explanation is that ZCWPW1 is a 'self-guarding' ISG, where the primary anti-viral function is self-guarded by a secondary IFN-repressing function. An example of this is MORC3, which restricts HSV-1 directly. When degraded by an HSV-1 virulence factor or knockdown by siRNA, its IFN-repressing function is lifted, triggering IFN/ISG induction [299]. Repeating the siRNA knockdown experiments in IFN-treated cells can also help confirm the observed phenotype and help determine the importance of ZCWPW1 in the IFN-induced response towards PRCV and SARS-CoV-2. More interestingly, proteins from the MORC (1-4) and ZCWPW (1 and 2) families, along with lysine demethylase 1B (KDM1B), are the only proteins in humans and pigs that possess the zinc finger CW-Type domain. Meanwhile, suppression of another lysine demethylase KDM5B was shown to increase the expression of IFN-β and other inflammatory cytokines following virus infection [300]. All this information sheds light on the potential mechanism of action of ZCWPW1.

The screens also identified ISGs such as USF1, HOXD13 and NRL as potential zoonotic barriers against coronaviruses, as they restrict SARS-CoV-2 to a larger extent than PRCV (**Fig4.12B and Table4.13**). While initial validation confirmed the antiviral properties of USF1 and HOXD13, further characterisation suggests the differential potency might simply be attributed to differential replication kinetics between the two viruses in the first place (**Fig4.16C and D**). The dynamic range observed in the SARS-CoV-2 screens is much wider than in the PRCV screens, meaning the most potent hit (i.e. IRF1) can restrict SARS-CoV-2 by 1000-fold (i.e. -10 Log2 fold change) but less than 50-fold (i.e. -5 log2 fold change) for PRCV (**Fig4.10A**). Apart from SARS-CoV-2 being able to replicate to a higher titre compared to PRCV, the fact that PRCV is more resistant

to porcine IFN response in NSK cells might also contribute to this (**Fig3.10B**). The difference in the dynamic range has made it difficult to directly compare the effect of individual ISGs between the two viruses.

Although they may not be cross-species restriction factors, validation results clearly showed that these ISGs have antiviral activities against both PRCV and SARS-CoV-2 (**Fig4.15C and D**). The fact that these ISGs were also identified in separate screening experiments against IAV (the same screening experiments which identified PHACTR1 and PHACTR4) suggests they may have a broad antiviral activity. Both of these ISGs have transcription factor activity, and therefore rather than acting directly, it is likely they exert their antiviral functions through activating transcription of other genes. For example, the relatively novel ELF1 identified from the PRCV screens (**Fig4.6 and Table4.7**) is also a transcription factor, and a recent report has shown that it initiates a broadly antiviral program by upregulating set of genes distinct from the type I IFN response [292].

While both USF1 and HOXD13 are highly conserved between humans and pigs ($> 95\%$), to the best of our knowledge, they have not been described to be antiviral in any human studies. One possible explanation is that neither gene was considered an ISG in human, as their upregulation following IFN response is either very subtle or not detected [154, 301], and therefore they have not been included in previous human libraries. Another possible explanation is that, despite the transcription factors are highly conserved, their binding sites may not be conserved. The region they bind to in pigs and humans may be different and potentially trigger the transcription of an entirely different set of genes.

Due to time constraints, we did not validate all candidates from the list of potential cross-

species coronavirus restriction factors **(Table 4.13)**. However, other candidates which are not transcription factors would potentially have a higher chance of being genuine cross-species restriction factors. This is because their antiviral function will likely be exerted via direct interaction with viral components, and therefore more likely to be counteracted by specific viral factors. Moreover, as mentioned, transcription factors are generally highly conserved among species, and therefore viral factors are more likely to be able to block the same transcription factors in different species. Initial validation for the candidates, similar to what was done for the candidate pan-coronavirus restriction factors **(Table 4.11)**, should be carried out to avoid false negatives.

As discussed, the ISG expression profile between cell lines can be different, evidenced by the differential sensitivity of PRCV towards porcine IFN in ST and NSK cells (**Fig3.8C and D**). Unlike the human ISG library, which was constructed based on multiple independent studies on multiple cell types and tissues. Our porcine ISG library is largely based on one recent publication measuring the ISG expression profile in one cell type [154]. The relative lack of studies and databases such as Interferome (http://www.interferome.org) for porcine ISGs has hindered the generation of a more comprehensive and unbiased library. Moreover, the human reference genome is better annotated and more accurate than the porcine reference genome. Our library is constructed based on a previous version of the reference genome (Sscrofa10) and the coding sequence for some of the ISGs in our library may not be up to date. The library may have little overlapping ISGs with the actual ISG profile of NSK cells, on which the screens were based, limiting the potential to identify IFN effectors responsible for the differential sensitivity between PRCV and SARS-CoV-2 towards porcine IFN specifically observed in NSK-derived cells

(**Fig3.10B**). A thorough comparison between the library and the latest reference genome to replace outdated sequences, along with repeating the RNA-sequencing experiment on more porcine cell types and tissues while generating our own database in the future can enhance the reliability and quality of our library for future experiments.

The differential expression profiles between cell lines can also directly affect the outcome of an experiment. For example, genome-wide CRISPR screens against SARS-CoV-2 in Vero, Huh7.5, and A549 cells from the same study were reported to have only two overlapping hits [302]. Unlike loss-of-function screens where a gene needs to be expressed in the first place for knockdown or knockout, gain-of-function screens merely use the cell as a vessel for overexpression. The cell type used should have less of an impact unless the function of the ISG depends on interaction with other cellular factors. Some ISGs do not work alone and function in pairs or larger complexes. For example, cGAS requires STING to phosphorylate IRF3, RNase L activation depends on the expression of OAS proteins, while ZAP antiviral function requires TRIM25 and the nuclease KHNYN [288, 303–305]. An antiviral ISG will have no or reduced effect when overexpressed in cells that lack the factors required by that specific ISG. Indeed, a study using a mass-spectrometry-based survey identified numerous interactions between ISGs and other ISGs or non-ISGs, with nearly 90% of them never reported [306]. This further suggests the majority of ISGs exert their antiviral potency via interaction with one or more proteins, and the nature of overexpression screens limited their use to address these questions. We initially wanted to carry out the screens in NPTr cells as they are a more relevant cell type for respiratory coronaviruses as NSK cells might have a different expression profile, lacking certain critical co-factors required for some ISGs in our library to function

against coronaviruses. Repeating the screen in different cell lines may uncover tissue or cell-type specific antiviral ISGs. Validating the hits in more cell lines may also provide more insight into their antiviral mechanisms.

It was initially surprising that other well-defined direct-acting ISGs such as IFITMs and Mx were not identified in the screens. During optimisation of the screens, it was found that certain lentiviruses, especially those encoding a potent antiviral ISG such as IRF1, IRF7, MDA5 and other components of the IFN pathway, generally have a lower RFP fluorescence intensity (see **Fig4.10C**). Nonetheless, they were able to produce a strong phenotype in transduced cells, restricting both PRCV and SARS-CoV-2 replication. Two hypotheses were made in an attempt to explain this. First, these ISGs may have a lower effective dose, meaning a subtle increase in expression can trigger a cascade of IFN response to restrict virus replication, and therefore the phenotypes they produce do not correspond to their transduction efficiency (i.e. RFP intensity). This can also potentially explain why direct-acting ISGs do not exhibit strong phenotypes in the screens, as they may be more dose-dependent. ISGs that trigger the IFN response may also impact lentivirus production, resulting in lower lentivirus titres. Second, these ISGs may have a broader effect on the cells, reducing the expression of RFP by affecting transcription or translation. Even if the cells were transduced with a high titre of lentivirus, the RFP signal they produce will be limited. In this case, the RFP intensity becomes a poor surrogate for measuring transgene expression level. Nonetheless, measuring the RFP fluorescence intensity of transduced cells remains an easy and rapid way to determine efficiency, especially in a medium-to-high throughput screening experiment. Alternative approaches such as RT-qPCR or western blot might be better for direct measurement

of transgene expression when characterising individual hits, but practically challenging.

The overall correlation between duplicate screens compared to other published studies is low due to the nature of the direct lysis RT-qPCR. Potential impurity in lysates can lead to greater variation compared to RT-qPCR using purified RNA (**Fig4.1C**). ISGs that do not have an effect on coronavirus replication can potentially give rise to CT anywhere between $\pm$ 1 from the GFP control. These "non-hits" comprise most of the ISGs in the screen and have negatively impacted the R^2 between duplicates. In fact, if only the hits from the SARS-CoV-2 screen were taken into consideration (**shaded area of Fig4.8B**), the R^2 would be 0.6, suggesting a relatively strong correlation. For the PRCV screen, the significant difference in transduction efficiency has contributed to the relatively weak correlation between the two replicates.

CHAPTER 5

SARS-CoV-2 replication in a hACE2 transgenic pig model

5.1 Overview and Background

In vitro models such as cell lines and primary cells are commonly used for basic research of respiratory infections as they are readily available, affordable and easy to maintain. However, their resemblance to *in vivo* infection is limited. Other *ex vivo* models such as air-liquid interphase culture and organoid have higher relevance and can recapitulate some cytopathology and inflammatory responses following infection. Nonetheless, the use of animal models is necessary for testing hypotheses about the mechanisms of diseases *in vivo*, evaluating antiviral therapeutics and vaccines and ensuring their safety before starting clinical trials in humans.

The primary animal species used to study SARS-CoV-2 infection and disease are mice, hamsters and non-human primates (NHPs). Mice are easy to manipulate, relatively cheap, and have wide availability of reagents. For these reasons, they are the most popular model and have been tested for susceptibility towards SARS-CoV-2 early in the pandemic. However, it was soon found that the spike protein of SARS-CoV-2 cannot effectively bind mouse ACE2, and strategies have been developed to solve this problem [307]. For example, mouse-adapted SARS-CoV-2 can be generated by sequential passaging in mouse lung tissue or by reverse genetics to modify the RBD [308, 309]. However, mouse-adapted SARS-CoV-2 viruses significantly differ from SARS-CoV-2 viruses circulating in the human population, especially in the RBD, limiting their use in studying neutralizing antibody responses [310]. Another approach is to generate transgenic animals expressing human ACE2, such as K18-hACE2 mice. This model was developed in response to the 2003 SARS-CoV epidemic and expresses human ACE2 under

the control of the human cytokeratin 18 promoter [227]. The K18-hACE2 mouse model is susceptible to most SARS-CoV-2 variants and shares some clinical characteristics observed in COVID-19 patients, such as upregulation of proinflammatory cytokines and chemokines, lung dysfunction, signs of dyspnoea, and anosmia [310, 311]. However, they may exhibit more severe disease and symptoms less confined to the respiratory system than that of humans, as virus is frequently detected in the brains of infected K18 mice [312, 313]. In general, SARS-CoV-2 infected mice do not accurately replicate the pattern of disease seen in humans, and the pattern of infection and immunological responses to SARS-CoV2 are significantly different from humans.

Another popular rodent model for SARS-CoV-2 is the Syrian golden hamster. Unlike mice, their ACE2 is compatible with the RBD of SARS-CoV-2 spike proteins and is therefore naturally susceptible. Compared with humans, hamsters show mild-to-moderate disease with fewer clinical signs after a very short incubation period (i.e. 1 to 2 days post-infection). They do not develop fever and generally fully recover by 2 to 3 weeks after infection [310, 311, 314]. SARS-CoV-2 infection in golden hamsters generally does not result in lethal outcomes, and mortality is only reported when challenged with high doses [315]. Diffuse alveolar damage (DAD), a hallmark of severe COVID-19, is observed in hamsters, but it is not associated with increased morbidity or mortality [311]. Interestingly, hamster models share some demographic differences observed in humans, where males and older individuals seem to develop more severe disease [316, 317]. Overall, the hamster model mimics the clinical, virological, histopathological and immunological features of mild-to-moderate disease in humans [314]. Therefore, there is an increasing use of hamsters over other animals to study the pathogenesis of SARS-

CoV-2, especially in the upper respiratory tract. However, similar to mice, hamsters are small animal models and certain important clinical markers and biopsies are difficult to measure and harvest. Moreover, the availability of hamster-specific research tools is scarce, limiting the use of this animal model.

Limitations of small animal models have led to the use of larger models such as NHPs as they exhibit the highest degree of genetic, anatomic, physiologic, and immunologic similarity to humans. Macqaue, marmosets and baboons are naturally susceptible to SARS-CoV-2 and have been used to recapitulate mild-to-moderate COVID-19 respiratory disease. Like humans, SARS-CoV-2 infection in NHPs is highly variable among individuals, with most cases resulting in asymptomatic infection while severe clinical disease and acute respiratory distress syndrome are rare. Similar to hamsters, older NHPs present more severe pulmonary disease, with higher viral burden and prolonged virus shedding from the nose and throat [318, 319]. However, there is no sex-associated difference in disease severity. SARS-CoV-2 infection in NHPs induces innate, humoral and cellular immune responses, which provide natural protective immunity against rechallenge [320, 321]. The Pfizer–BioNTech vaccine and the antiviral drug remdesivir were first tested in NHPs before the start of human clinical trials, demonstrating their importance [322, 323]. Nonetheless, the use of NHPs is limited by the scarcity of required expertise and facilities to handle them. The high cost, low supply and related ethical issues of NHPs have also limited the sample size in individual studies, which may not be sufficiently powered to detect differences between experimental groups.

Despite substantial immunity through vaccination and previous exposure, emerging vari-

ants of SARS-CoV-2 continue to spread rapidly, causing morbidity and mortality. The use of animal models to test new therapeutic agents and vaccines remains necessary. Existing models each have their pros and cons as discussed, and developing alternative models to fill in the gaps can allow a better understanding of COVID-19-associated disease and potential medical countermeasures. Apart from NHPs, there is a lack of large animal models and models that can accurately reflect the pathology associated with severe COVID-19 in humans. Meanwhile, the success rate of translating therapeutic measures developed in small animal models to humans is very low. These have urged the development of alternative large animal models with similar physiology to humans to act as intermediate models between rodents and NHPs.

There is an increasing appreciation of livestock as biomedical models, with pigs being one of the most popular. Pigs are genetically, anatomically, physiologically and immunologically closer to humans than many existing animal models [218]. In fact, pigs are the most commonly used species in xenotransplantation due to the similarity of their organs to human organs. Compared to NHPs, pigs are relatively inexpensive, accessible and more ethically acceptable. Although the lungs of humans and pigs have slight anatomical differences (porcine lungs have two lobes on the left and four lobes on the right, whereas humans have only three on the right), their overall size and development are similar. Pigs are therefore used extensively for the study of respiratory diseases such as influenza and tuberculosis. Early attempts were made in the COVID-19 pandemic to test if pigs are suitable models, given they have been previously reported to be susceptible to SARS-CoV-1 with *in silico* data also suggesting porcine ACE2 (pACE2) can bind the spike protein of SARS-CoV-2 [310, 324]. However, several studies carried out by different

research groups indicated that pigs are not naturally susceptible to SARS-CoV-2 and no viral replication was detected [325, 326]. Other studies utilizing primary cells from respiratory tracts, *ex vivo* air-liquid interface cultures and precision-cut lung slices from pigs have yielded similar results [327, 328]. Interestingly, as shown in previous chapters and reports from other groups, certain porcine cell lines are highly susceptible to SARS-CoV-2 *in vitro*.

The lack of susceptibility of pigs towards SARS-CoV-2 infection has limited their use as a disease model. To address this, we have generated a transgenic pig model expressing hACE2 and subjected it to an *in vivo* challenge study using an early isolate of SARS-CoV-2. This project requires expertise in large animal handling, immunohistochemistry, and pathology, and it would not have been possible without the collaboration of the Griffiths group at Moredun Research Institute (hereafter Moredun) and other departments within the university. All *in vitro* data were generated by myself unless otherwise specified. Other *in vivo* experiments and histology staining and scoring were carried out by the specified collaborators as stated in detail in the result section.

5.2 Results

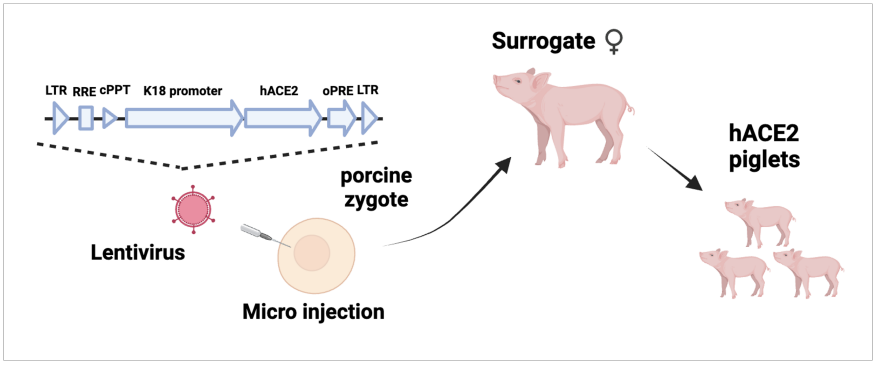
5.2.1 Generation of hACE2 transgenic pigs and *in vitro* characterisation

In previous chapters, we have hypothesised the difference in IFN response between humans and pigs might contribute to barriers against *in vivo* SARS-CoV-2 infection in pigs. Nonetheless, the likeliest determinant remains to be receptor compatibility and expression, where pACE2 might not be sufficiently expressed in the correct tissue or cell type for SARS-CoV-2 utilisation *in vivo*. To address this, we have generated a transgenic pig model expressing hACE2 and tested if this is sufficient to render pigs susceptible to SARS-CoV-2 infection.

Previous attempts by another group to generate susceptible pigs were made by introducing hACE2 to the locus of endogenous pACE2 by CRISPR/Cas9 technology [329]. However, no *in vivo* challenge studies were described following the initial report. We suspect the attempt was unsuccessful and have adopted an alternate strategy. The K18-hACE2 mice mentioned above were generated by direct injection of plasmid DNA expressing hACE2 under the human keratin 18 (K18) promoter and have proven to be a useful model [227]. However, the chance for random integration of plasmid DNA into genomic DNA is low.

To generate the transgenic pigs, a lentivirus expressing hACE2 under the K18 promoter was adopted. Unlike the CMV promoter (used in previous chapters), which is non-specific

A



B

Dam	Pig	Sex	Proviral (CT)
5386	P13	F	19.73
5386	P7	F	20.86
5386	P1	M	20.88
5386	P2	M	21.07
5258	P20	F	21.2
5258	P17	M	21.23
5258	P26	F	21.37
5386	P10	F	21.5
5258	P22	F	21.63
5386	P9	F	21.85
5258	P21	F	21.95
5258	P18	M	22.05
5386	P8	F	22.16
5386	P5	F	22.18
5386	P4	M	22.34
5305	P30	M	22.36
5258	P15	M	22.38
5386	P12	F	22.62
5386	P3	M	22.79
5258	P14	M	22.85
5258	P19	M	23.62
5386	P6	F	23.88
5258	P16	M	25.46
5258	P27	F	27.1
5258	P24	F	27.15
5258	P25	F	27.8
5305	P29	M	27.81
5305	P28	M	29.75
5386	P11	F	30.76
5258	P23	F	ND
5305	P31	F	ND
5305	P32	F	ND

C

Parentage	Cohort	Sex	Pig	Relative hACE2 expression
P2xP13	2	F	P61	53.82
P2xP7	2	F	P46	51.27
P2xP13	1	F	P59	39.95
P2xP7	1	F	P48	31.12
P2xP13	2	F	P60	27.28
P2xP13	1	F	P57	25.99
P2xP7	3	F	P43	25.46
P2xP13	3	M	P52	19.56
P2xP13	3	M	P53	19.16
P2xP13	UN	F	P56	15.45
P2xP13	UN	F	P58	15.45
P2xP13	X	M	P54	15.35
P2xP13	X	M	P51	13.18
P2xP7	X	M	P39	8.88
P2xP7	X	F	P44	7.46
P2xP7	X	M	P37	7.21
P2xP13	X	M	P55	6.15
P4xP26	X	F	P62	2.39
P2xP7	X	F	P50	2.23
P2xP7	X	F	P45	1.71
P2xP7	X	F	P47	1.49
P2xP7	X	M	P34	1.29
P2xP7	X	M	P36	1.06
P2xP7	X	M	P33	1
P2xP7	X	F	P42	0.23
P2xP7	X	F	P49	0.03
P2xP7	X	M	P40	0.03
P2xP7	UN	M	P38	0.02
P2xP7	UN	M	P35	0.02

Figure 5.1: **Generation of hACE2 transgenic pig**. **(A)** Schematic representation of the generation of hACE2 transgenic pigs. Briefly, a lentivirus expressing human ACE2 under the tissue-specific keratin 18 promoter was microinjected into porcine zygotes and implanted into surrogates (created with Biorender.com). **(B)** Relative levels of lentivirus integrated into the genome of F0 transgenic pigs were determined using the Lenti-X Provirus Quantitation Kit. The surrogate dam and sex (M=male, F=female) of each pig are indicated. Pigs selected for breeding of F1 animals are highlighted in green (female) and blue (male). Texts in grey indicate pigs that were culled due to bacterial infection. ND= not detected. **(C)** Total RNA extracted from ear biopsies of F1 piglets were subjected to RT-qPCR to determine human ACE2 transcript levels. Expression levels were normalised to GAPDH and presented relative to P33. Parentage is shown and the allocated cohort for the challenge study is indicated for each pig. UN indicates pigs used as uninfected controls. Texts in grey indicate pigs that were not used for the challenge study. Pigs are ranked based on human ACE2 proviral or expression level.

and will promote transgene expression in many cell types, the K18 promoter is cell-type specific, allowing efficient transgene expression in epithelial cells lining the liver, kidney, gastroenteritis tracts and most importantly, the site of SARS-CoV-2 infection [227]. The lentivirus was microinjected into the perivitelline space of putative zygotes, which were then surgically implanted into five surrogate gilts. Three were confirmed pregnant and a total of 32 hACE2 transgenic founder piglets (F0) were produced **(Fig5.1A)**. Genomic DNA was extracted from the ear biopsies collected from these F0 piglets, and the levels of lentiviral integration in each animal were determined by the Lenti-X Provirus Quantitation Kit (Takara) (**Table5.1B**). Three gilts and two boars with the highest proviral levels were selected for breeding and generated 29 second-generation transgenic animals (F1). Total RNA was extracted from their ear biopsies and subjected to RT-qPCR to determine the relative hACE2 transcript level in the F1 piglets **(Table5.1C)**. The cloning and generation of the lentivirus stocks was performed by Dr. Finn Grey. The generation and breeding of animals were carried out by trained professionals based at the large animal research and imaging facility (LARIF) and Dr. Simon Lillico, a senior post-doctoral fellow at the Roslin Institute.

Prior to *in vivo* challenge, the *in vitro* properties of cells from the transgenic pigs were characterised. Primary fibroblast cells were generated from the ear biopsies of the nine F1 piglets with the highest hACE2 expression levels. However, samples from P57 were contaminated by bacteria and no cells were recovered. Two additional biopsies from pigs with low transcript levels were included as controls (P33 and P38). A total of ten cell lines were generated and their total RNA was extracted, with hACE2 expression levels determined by RT-qPCR (**Fig5.2A**). The expression levels of the generated fibroblast

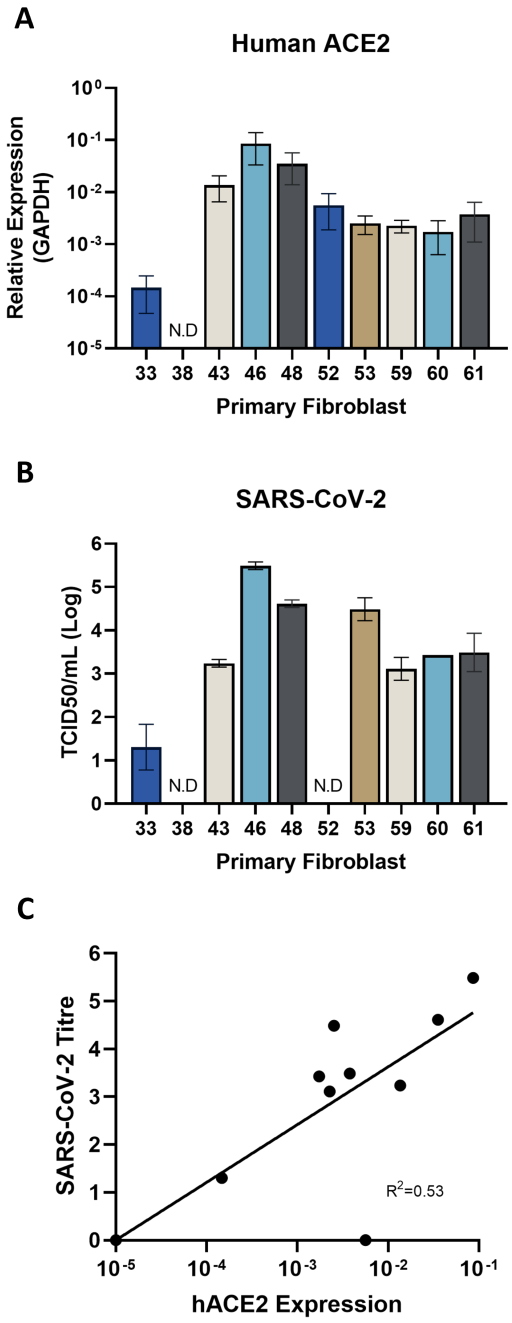


Figure 5.2: **hACE2 transcript levels correlate with susceptibility to SARS-CoV-2** Primary fibroblast cells were generated from the ear biopsies of F1 transgenic pigs. **(A)** Total RNA was extracted and subjected to RT-qPCR to determine the hACE2 transcript level. Expression levels were normalized to porcine GAPDH. **(B)** Cells were seeded in 24-well plates and infected by SARS-CoV-2 isolate EDB2 at an MOI of 3. Supernatant samples were harvested 24 HPI and titred by TCID50 on hACE2-NSK cells. **(C)** Correlation between hACE2 expression and susceptibility towards SARS-CoV-2 in the primary fibroblast cells. The coefficient of determination(R^2) was calculated from a simple linear regression. For calculation, an arbitrary expression value was assigned to cells from P38 as hACE2 was undetectable. Results represent the average of two independent experiments, each with two technical replicates. Error bars represent the standard deviation.

cells are generally consistent with the expression levels observed in their respective biopsies (**Table5.1C**). The hACE2 transcript levels in cells from P33 and P38 were low and undetectable respectively. Expression levels were the highest in cells from P46 and P48, which are among the highest in the biopsied samples.

To determine the susceptibility of the primary fibroblasts towards SARS-CoV-2, cells were seeded in 24-well plates and infected by EDB2 at an MOI of 3. Supernatant samples were harvested 24 HPI and titred by TCID50 on hACE2-NSK cells. As expected, cells from P46 and P48 with the highest hACE2 expression levels produced the highest EDB2 titre. Extensive CPE was also observed in these cells (data not shown). In contrast, very low levels of virus replication were detected in cells from P33 and no virus was detected from cells of P38. Surprisingly, despite the relatively high level of hACE2 expression, cells from P52 do not seem to be susceptible to EDB2 (**Fig5.2B**). Otherwise, there is a strong correlation between hACE2 transcript level and *in vitro* susceptibility towards EDB2 in the porcine primary fibroblast cells (**Fig5.2C**).

Apart from EDB2, the susceptibility of the isolated primary fibroblast was also subjected to infection of two additional SARS-CoV-2 variants, Delta (Lineage B.1.617.2) and Omicron. Unlike EDB2, neither the Delta nor Omicron variants produced CPE on the cell line (i.e. hACE2-NSK) we employed for TCID50 titring. Alternatively, the SARS-CoV-2 N copy number in the supernatant was measured by RT-qPCR instead.

Similar to the previous experiment, cells were seeded in 24-well plates and infected by EDB2, Delta or Omicron variant at an MOI of 3. Supernatant samples were harvested 48 instead of 24 HPI in order to capture the slower replication kinetics of the Delta and

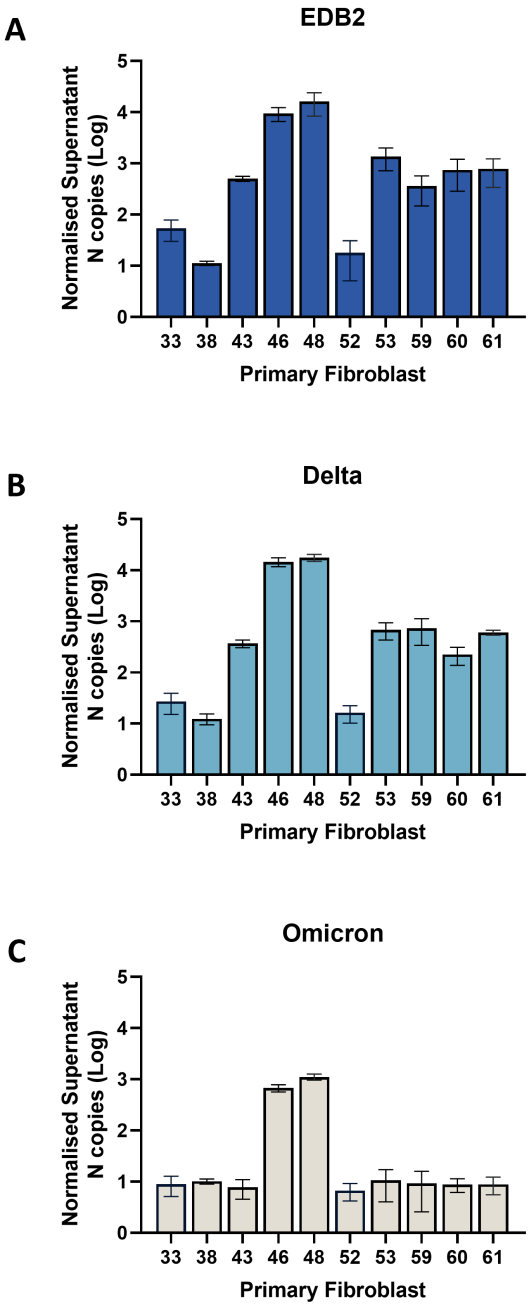


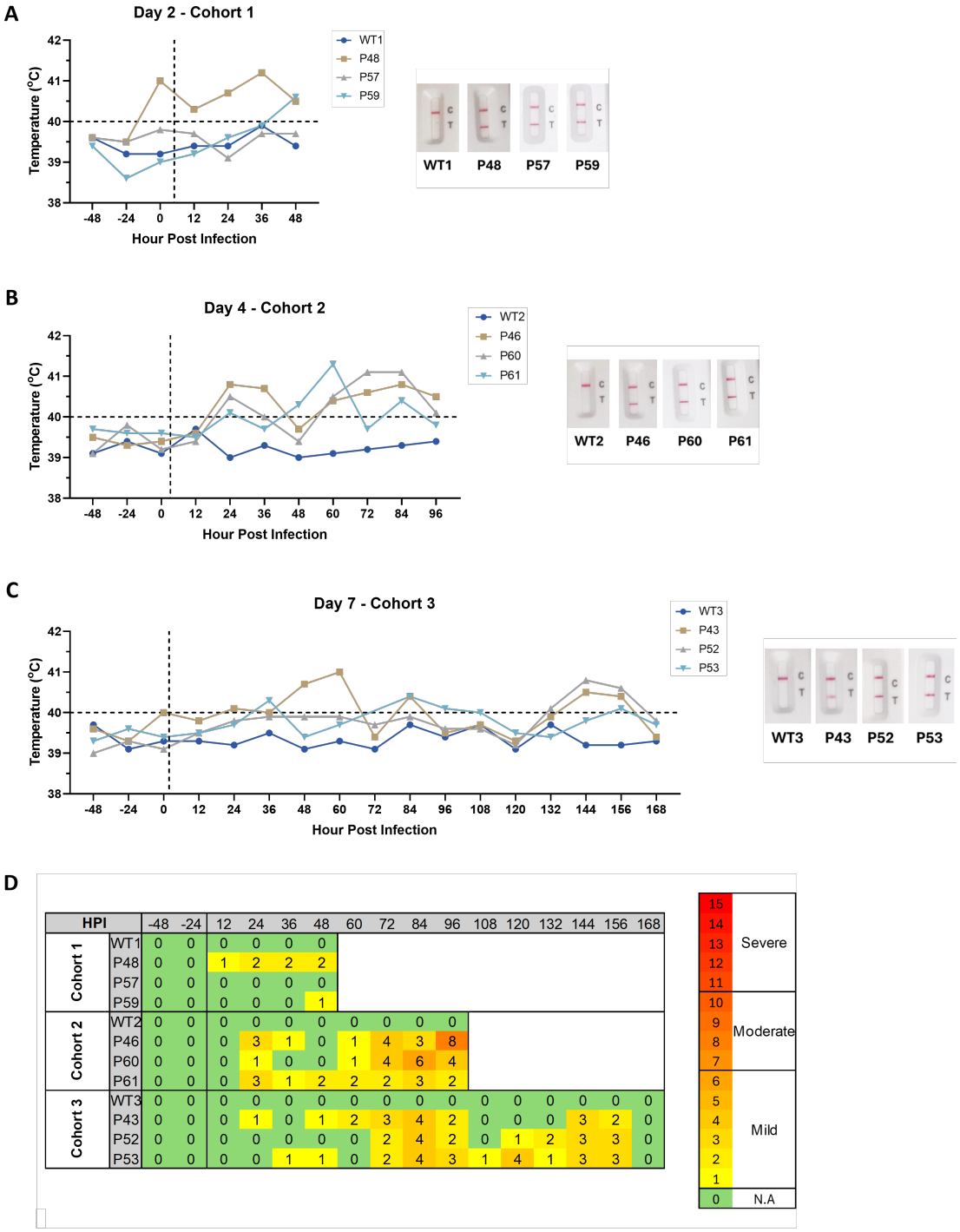
Figure 5.3: **Susceptibility of primary fibroblast cells isolated from transgenic pigs to other SARS-CoV-2 variants** Primary fibroblast cells were generated from the ear biopsies of F1 transgenic pigs. Cells were seeded in 24-well plates and infected by SARS-CoV-2 isolate **(A)** EDB2, **(B)** Delta or **(C)** Omicron variant at an MOI of 3. Supernatant samples were harvested 2 and 48 HPI and subjected to Trizol-LS RNA extraction. The level of the nucleocapsid (N) gene is measured by RT-qPCR. Results represent the average of three independent experiments, each with two technical replicates. Error bars represent the standard deviation.

Omicron variants. Unlike EDB2, where extensive CPE was observed in cells with high levels of hACE2 (i.e. P46 and P48), none of the cells infected by the other two variants produce any form of CPE. The copy number data for EDB2 **(Fig5.3A)** is consistent with the TCID50 data **(Fig5.2B)**, suggesting the measurement of N copies is representative of infectious virus particle released. Similar to EDB2, the *in vitro* susceptibility of the primary fibroblast cells to the Delta variant correlates with hACE2 expression, where a high level of N is detected from P46 and P48 while P33 and P38 produced the lowest level of N copies **(Fig5.3B)**. In contrast, this correlation was not observed in the Omicron variant. Only cells from P46 and P48, with the highest hACE2 expression, are susceptible **(Fig5.3C)**. Consistent with previous experiments, cells from P52 are not susceptible to any SARS-CoV-2 variants despite the relatively high level of hACE2 expression.

5.2.2 *In vivo* SARS-CoV-2 challenge of transgenic pigs

The primary cells isolated from transgenic animals were confirmed to be susceptible to SARS-CoV-2 infection. We have therefore selected nine F1 animals with the highest transgene expression for the *in vivo* challenge study. The nine trangenic pigs, along with three genetically similar wild-type control pigs were inoculated with 1 x 10^6 TCID50 of EDB2 intranasally. They were initially assigned into three cohorts, each with three transgenic and one wild-type animal, based on hACE2 expression levels from biopsy samples (**Table5.1C**) to ensure even distribution of pigs expressing varying levels of hACE2, and designed to be euthanised at 2, 4 and 7 days post-infection (DPI). However, as the study progressed, the design was altered on the advice of the resident veterinarian, with pigs showing more severe clinical signs selected for earlier time points to avoid fatal outcomes from the viral challenge. The animal license that covered these studies, only extended to moderate outcomes, with fatalities being a notifiable event. Rectal temperatures and clinical status were recorded every 12 hours and SARS-CoV-2 lateral flow tests (LFT) were performed with nasal swabs collected post-mortem for each cohort (**Fig5.4**). Clinical assessment, euthanasia and necropsy of animals were carried out by veterinarians based at Moredun.

All infected transgenic animals show varying degrees of clinical symptoms including coughing, sneezing, loss of appetite and increased body temperature. Most importantly, all transgenic pigs tested positive by LFT, as early as 2 DPI. In contrast, none of the wild-type animals showed any sign of infection and were LFT negative, consistent with previous reports suggesting pigs are naturally non-susceptible to SARS-CoV-2. The



HPI		-48	-24	12	24	36	48	60	72	84	96	108	120	132	144	156	168
Cohort 1	WT1	0	0	0	0	0	0										
	P48	0	0	1	2	2	2										
	P57	0	0	0	0	0	0										
	P59	0	0	0	0	0	1										
Cohort 2	WT2	0	0	0	0	0	0	0	0	0	0						
	P46	0	0	0	3	1	0	1	4	3	8						
	P60	0	0	0	1	0	0	1	4	6	4						
	P61	0	0	0	3	1	2	2	2	3	2						
Cohort 3	WT3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P43	0	0	0	1	0	1	2	3	4	2	0	0	0	3	2	0
	P52	0	0	0	0	0	0	0	2	4	2	0	1	2	3	3	0
	P53	0	0	0	0	1	1	0	2	4	3	1	4	1	3	3	0

15, 14, 13, 12, 11	Severe
10, 9, 8, 7	Moderate
6, 5, 4, 3, 2, 1	Mild
0	N.A

Figure 5.4: **hACE2 transgenic pigs present clinical signs following SARS-CoV-2 infection.** Three cohorts comprising a total of nine transgenic and three wild-type pigs were inoculated with 1 x 10^6 TCID50 of SARS-CoV-2 isolate EDB2 intranasally. Rectal temperatures of infected animals were recorded every 12 hours and nasal swabs were collected post-mortem for lateral flow tests for **(A)** Cohort 1 (euthanised on Day 2) **(B)** Cohort 2 (euthanised on Day 4) and **(C)** Cohort 3 (euthanised on Day 7). **(D)** The Clinical status of infected animals was monitored twice a day. A clinical assessment scheme was established prior to the study, in which pigs were scored against several criteria including demeanour, appetite, temperature and other respiratory signs (cough, sneezing). Cumulative scores ≥ 7 are defined as moderate severity in this model. HPI = hours post-infection.

normal rectal temperature of pigs lies between 38.7–39.8°C, and 40°C or above is considered febrile [330]. P48 from the first cohort showed signs of fever and some clinical symptoms before the challenge began, suggesting it may have an irrelevant underlying health condition. Otherwise, pigs in this cohort did not develop obvious symptoms before euthanisation. In contrast, transgenic animals from the second and third cohorts began developing symptoms as early as 1 DPI, and severity peaked on day 3 or 4. The clinical manifestation of infected transgenic animals matched that of mild-to-moderate COVID-19 patients, suggesting they are susceptible to SARS-CoV-2.

Tissues, including the nasal turbinate, trachea and lung were collected from euthanised animals and analysed for transgene expression, virus detection and histological analysis. Immunohistochemistry was performed by the histology team based at the Easter Bush pathology department. Immunohistochemical staining of formalin-fixed tissue samples showed varying expression of hACE2 protein in the respiratory tracts of transgenic pigs, with representative images from P46 shown (**Fig5.5B**). To objectively quantify the hACE2 staining signals from all samples, the Qupath software for digital pathology was implemented to calculate the pixel-wise H-score (see methods for detail). This allows a direct comparison of hACE2 expression in the different tissues of all pigs. No hACE2 was detected in any tissue in the wild-type animals.

RT-qPCR revealed high levels of SARS-CoV-2 N copies throughout the respiratory tracts of transgenic pigs, especially in the nasal turbinates, at 2 and 4 DPI, ranging from $3.0x10^6$ to $5.0x10^7$ copies/µg total RNA. Levels in the trachea and lung were lower, ranging between $5.9x10^4$ and $5.9x10^6$ copies/µg total RNA for trachea and $1.3x10^3$ and

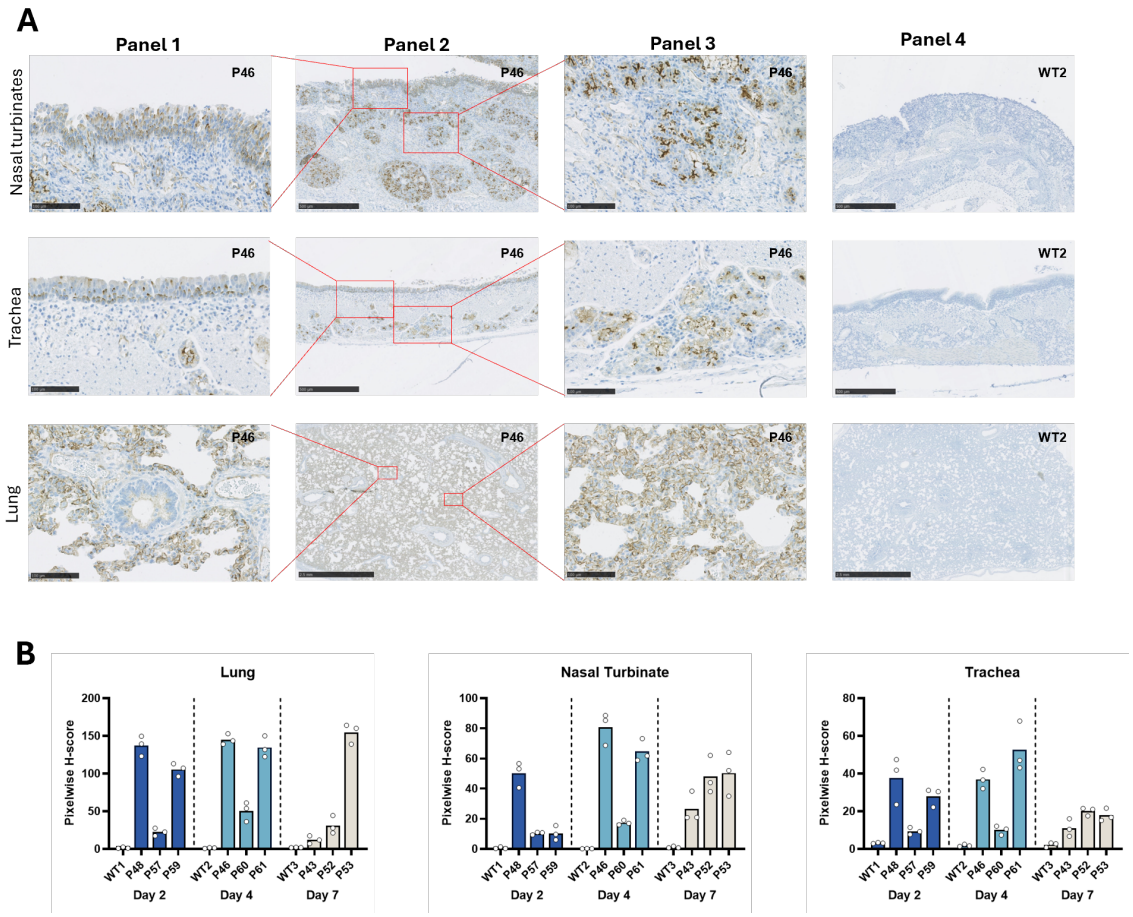


Figure 5.5: **Expression and distribution of hACE2 in pig tissues. (A)** Immunohistochemical staining demonstrating distribution of hACE2 expression in sections of nasal turbinate, trachea, and lung from P46 culled four DPI. Red boxes in panel 2 indicate the location of magnified images displayed in panels 1 and 3. Panel 4 shows the control WT pigs without hACE2 expression. Staining indicates hACE2 expression in ciliated columnar epithelial cells, goblet cells and seromucinous glands within the nasal turbinates and trachea. In lungs, staining is most pronounced in alveolar pneumocytes but to a lesser extent in bronchiolar epithelium and vascular endothelium. Scale bars for panels 1 and 3 are 100mm; for panels 2 and 4 are 500mm, apart from lung images which are 2.5mm. **(B)** Total hACE2 staining was quantified using QuPath software. Viable tissue regions on each whole-slide IHC image were randomly sampled three times by non-overlapping regions with areas no smaller than 4 mm^2. The pixel-wise H-score for each region was calculated (see Methods) and the average score was presented for each image.

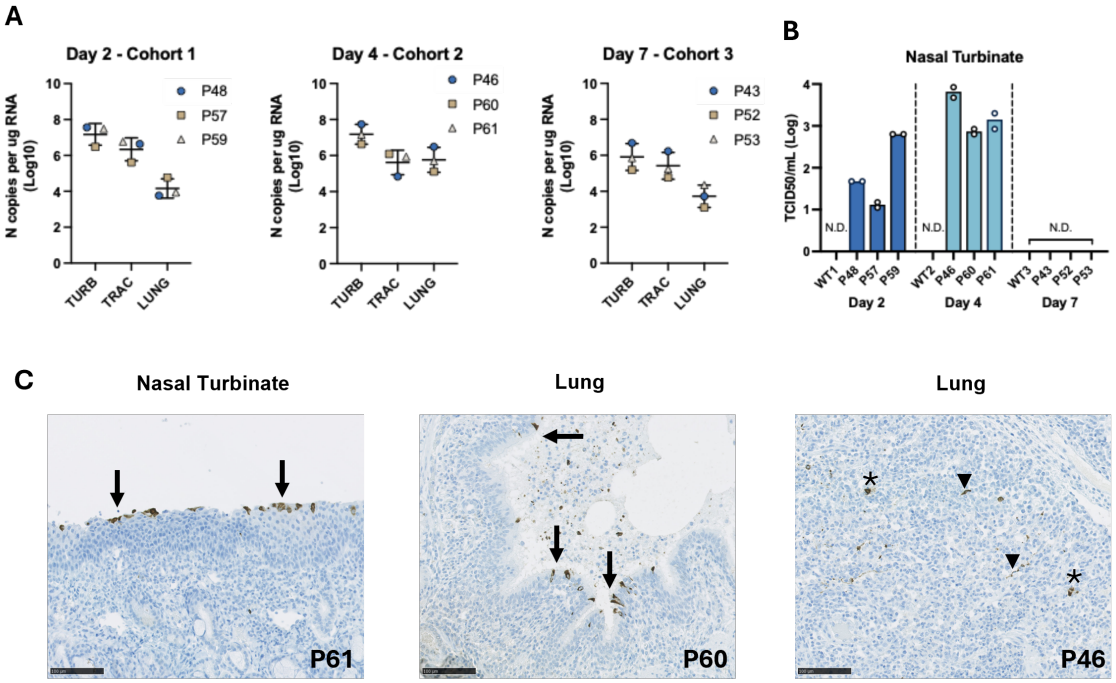


Figure 5.6: **SARS-CoV-2 RNA, protein and infectious virus were detected in respiratory tissues of infected transgenic animals.** Tissues were collected post-mortem from all animals used in the challenge study. **(A)** Samples were lysed in Trizol and total RNA was extracted for RT-qPCR. The copy number of SARS-CoV-2 N gene in the nasal turbinate (TURB), trachea (TRAC) and lung (LUNG) of transgenic pigs was determined. Each data point represents the average from 2 technical replicates. **(B)** Tissue samples were homogenised in serum-free DMEM and titred by TCID50 using hACE2-NSK cells. No virus was recovered from the trachea and lung samples. Data represents the mean from two independent experiments (each using two technical replicates). **(C)** Tissue sections from nasal turbinates and lung were stained for SARS-CoV-2 Nucleocapsid protein demonstrating the presence of viral protein within mucosal epithelial cells of both the upper and lower respiratory tract (arrow) as well as alveolar macrophages (*) and pneumocytes (arrowhead). Representative images shown from hACE2 transgenic pigs. (20x magnification).

$3.0x10^6$ copies/µg total RNA for the lung **(Fig5.6A)**.

To determine if any infectious virus can be recovered, a separate portion of tissue samples was homogenized in serum-free media and titred by TCID50 using hACE2-NSK cells. Infectious SARS-CoV-2 was recovered from the nasal turbinate but not from the trachea and lung, possibly due to the lower levels of RNA detected in these sites. A higher titre of virus was recovered from Day 4 than from Day 2, although similar levels of viral RNA were detected in both cohorts. No infectious virus was recovered from Day 7, consistent with the lower level of viral RNA detected in this cohort **(Fig5.6B)**. No viral RNA, protein or infectious virus was detected from any tissues taken from the WT control pigs.

IHC staining identified SARS-CoV-2 N protein within the cytoplasm of respiratory epithelia of nasal turbinates **(Fig5.6C)**. Within the distal lung parenchyma, N protein was identified in bronchiolar epithelial cells, macrophages, alveolar pneumocytes and occasional vascular endothelial cells. No SARS-CoV-2 N protein was detected in WT pig tissues.

Histological examination of lung tissues from infected transgenic pigs revealed significant inflammation as early as 4 DPI, displaying key histopathological features similar to those described in fatal COVID-19. P46 from cohort 2 (**Fig5.7**) shows the most severe lung pathology, including significant neutrophil and macrophage-rich inflammation in the lungs. Oedema, focal, fibrin-rich intravascular thrombi and bronchopneumonia with clear signs of various damage suggestive of DAD were also observed[331]. Samples were taken from multiple regions, suggesting that inflammation within the lung was focal in

nature, as separate samples from the same lungs displayed disparate levels of pathology. Even within the same tissue section, areas of severe lung inflammation were observed adjacent to less affected regions, possibly due to uneven distribution of the viral inoculum. Histological examination was performed by two independent pathologists, Prof Tanja Opriessnig (Moredun) and Dr David Doward (Royal Infirmary of Edinburgh) blinded to the status of the pigs.

The hACE2 transgenic pigs presented here can potentially serve as a model for vaccine studies as efficacy in vaccine protection can be determined following challenge experiments. Characterising the immune response to SARS-CoV-2 infection in the transgenic pigs will provide valuable baseline data for future vaccine studies. However, due to the design of this study, the infected pigs were kept up to only 7 DPI and full characterisation of the adaptive immune response which usually occurs later during the infection course is not possible. Nonetheless, total immunoglobulin (Ig) levels were determined by enzyme-linked immunosorbent assay (ELISA) from the serum of infected animals. Although considered early for antibody production, a strong antibody response was detected in P53 which was culled 7 DPI. Some responses were also detected in P43 at 7 DPI and P60 at 4 DPI, although neither reached the score considered positive (S/P ratio >60%, see methods for details). None of the WT pigs were seroconverted.

The use of lentiviral vectors for generating transgenic animals will result in variable transgene expression. *In vitro* studies on primary fibroblast cells indicate that hACE2 levels correlate with susceptibility to SARS-CoV-2 replication **(Fig 5.2C)**. If the same concept applies *in vivo*, levels of hACE2 determined by ear biopsies may be used to

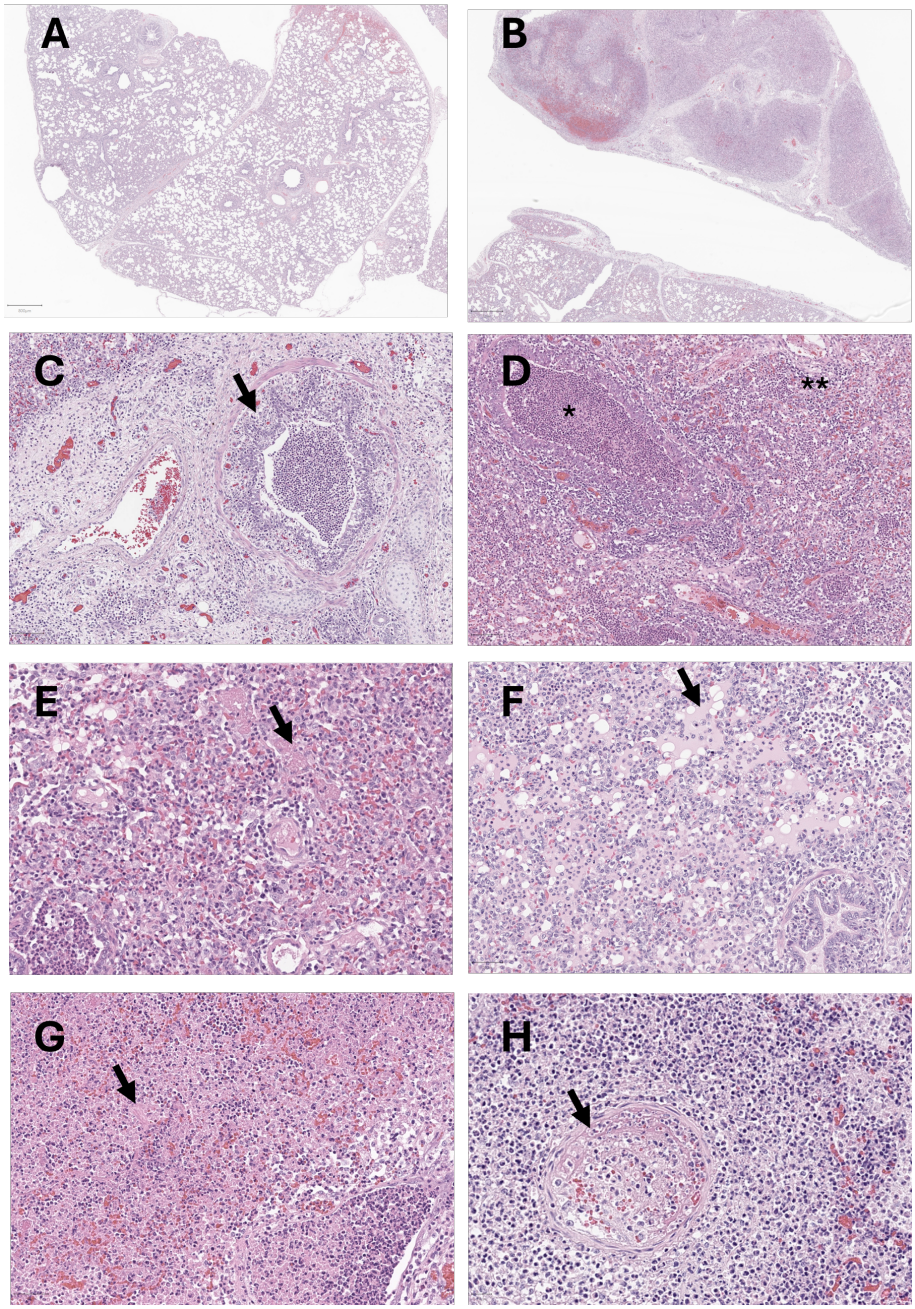


Figure 5.7: **SARS-CoV-2 induces severe inflammatory lung pathology in hACE2 transgenic pigs.** Lungs of **(A)** WT2 (non-transgenic infected) and **(B)** P46, both culled at four DPI. In P46, evidence of neutrophil-rich bronchial inflammation was observed. **(C)** Widespread bronchiolar (*) and alveolar (**) inflammation **(D)** with associated intra-alveolar fibrin **(E)**, oedema **(F)** and parenchymal necrosis **(G)**. Occasional organising (arrow) and fibrin-rich thrombi were present in medium calibre vessels **(H)**. Representative images were shown. **(A)** and **(B)** 1x magnification; **(C)** 10x magnification; **(D)** to **(H)** 20x magnification.

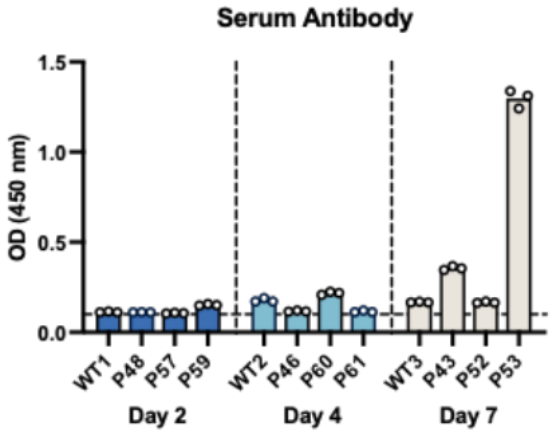


Figure 5.8: **Antibody response to SARS-CoV-2 infection in hACE2 pigs** Serum was generated from blood samples taken immediately after postmortem. Antibody levels were determined using a double antigen multi-species Elisa assay, recognizing SARS-CoV-2 N antigen. Despite relatively early for seroconversion, high levels of antibodies were detected in the serum of P53, with lower levels detected in P43 and P60. Data represents the mean from three technical repeats.

predict disease severity following infection with SARS-CoV-2 in individual animals. Using normalised rank scores, a heatmap ordered according to the average hACE2 pixel-wise H-score in the nasal turbinate, trachea and lung of transgenic pigs as shown in **Fig5.5B**, comparing to their respective SARS-CoV-2 levels and disease presentation was shown **(Fig5.9)**.

The analysis suggests a trend towards increased viral loads and more severe clinical and histological outcomes in pigs with higher levels of hACE2 expression. However, there are several caveats to this analysis. The relatively low numbers of animals used in the challenge study mean statistical significance cannot be achieved and variation in individual animals, unrelated to hACE2 expression, will confound potential correlation. Furthermore, scores will be biased depending on when the animals were culled. For example, peak viral loads occurred at four DPI. Higher levels of inflammation are also more likely to occur at later time points following accumulation of host responses to

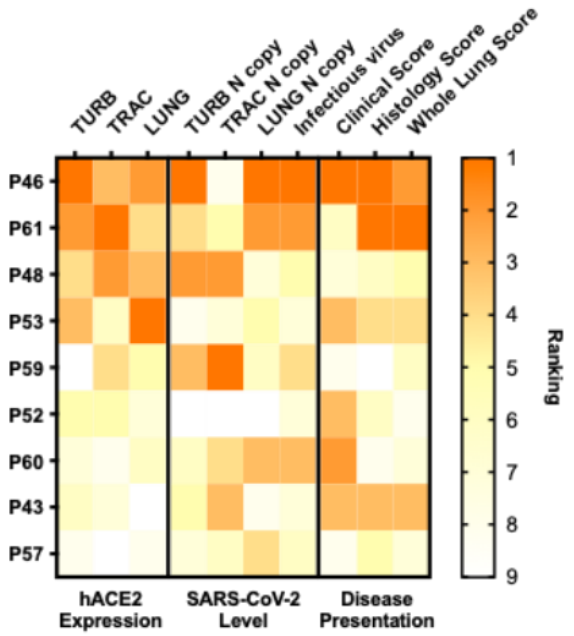


Figure 5.9: **Heatmap displaying correlation between hACE2 expression, SARS-CoV-2 virus level and disease presentation** Data were normalised by a rank-based system. Pigs ranked higher in a data category refer to a stronger phenotype or higher value in that specific category and vice versa. Note that pigs can be ranked equivalently in some data categories. The heatmap is ordered by the average pixel-wise H-score for hACE2 expression in the nasal turbinate (TURB), trachea (TRAC) and Lung (LUNG).

virus infection. Larger studies will be required to determine whether there is a statistical correlation between hACE2 expression levels and disease severity.

5.3 Discussion

Although existing animal models have been valuable in understanding COVID-19-related disease, they each have their pros and cons and there is a lack of easily accessible large animal models that can replicate the primary features of severe COVID-19: pulmonary inflammation. Pigs have been used extensively as animal models for respiratory disease due to their similarity to humans. However, they cannot be used in COVID-19-related research as they are not susceptible to SARS-CoV-2. To address this, we have generated a hACE2 transgenic pig model highly susceptible to SARS-CoV-2 with lung pathology that faithfully replicates severe COVID-19 in humans.

A previous report has also described the generation of a human ACE2 transgenic pig model in 2021 [329]. Consistent with our results, primary epithelial cells isolated from the lungs and kidneys of these animals were susceptible to SARS-CoV-2. However, no subsequent *in vivo* challenge studies were reported, and whether these pigs are susceptible to SARS-CoV-2 is unknown. In their study, the hACE2 cDNA was inserted at the pACE2 locus via CRISPR/Cas9-mediated homology-directed repair, meaning the transgene is expressed under the same endogenous promoter as pACE2. In contrast, we introduced hACE2 into the animals by lentiviral transduction, expressed under a separate promoter (i.e. K18) with the possibility of multiple copies inserted into multiple loci (explaining the differential proviral levels in the F0 pigs, **Table5.1B**). One of the possible explanations for SARS-CoV-2 (at least for EDB2) being able to utilise pACE2 for entry *in vitro* but not *in vivo* might be differential expression. The expression profile of ACE2 in pigs might differ from that of humans and may not be in the relevant cell types or

tissues for SARS-CoV-2 infection. In such cases, the expression of hACE2 under the endogenous promoter of pACE2 is insufficient to alter the physio-spatial expression of entry receptors. In contrast, the exogenous K18 promoter may have led to expression in the required tissues to enable susceptibility to SARS-CoV-2. Moreover, our transgenic animals likely resulted in a higher hACE2 expression, which might be critical for *in vivo* infections.

Indeed, a correlation between hACE2 expression levels and susceptibility towards SARS-CoV-2 was observed *in vitro*. As shown in **Fig5.2 and Fig 5.3**, primary fibroblast cells from transgenic pigs which express higher hACE2 transcripts produce a higher titre when infected by SARS-CoV-2 EDB2 and Delta. A similar trend was also observed in the case of Omicron, except only fibroblast cells with the highest hACE2 expression (i.e. P46 and P48) are susceptible. The only exception was cells from P52, which were not susceptible to all three variants for unknown reasons despite a high transcript level. This suggests susceptibility to SARS-CoV-2, especially to the Omicron variant was not solely dependent on the expression of hACE2, and other barriers likely exist. As discussed previously, compared to other VOCs, the early Omicron variant (i.e. BA.1 and BA.2) has an enhanced ability to evade adaptive immunity (i.e. antibody escape), but reduced ability to evade innate immunity [260, 261]. Later Omicron subvariants (i.e. BA.4 and BA.5) have however acquired the ability to evade innate immune suppression through convergent evolution [332]. The weaker innate immune suppression of BA.1 possibly accounts for its failure to infect most primary cells despite having substantial hACE2 expression for infection of EDB2 or Delta. This suggests, similar to the case of hACE2-NSK cells discussed in Chapter 3, that such barriers can be overcome by extensive

overexpression of entry rectors. It will also be interesting to investigate why fibroblast cells from P52 cannot support any SARS-CoV-2 replication, as the IFN response may potentially play a role as discussed in previous chapters.

The correlation is also, to some degree, observed *in vivo*. As shown in **Fig 5.9**, animals with higher hACE2 expression in their respiratory tracts (i.e. P46 and P61) tend to have higher viral levels, more severe clinical symptoms and lung pathology *in vivo* following SARS-CoV-2 infection. Also, P46 experienced the highest fever (40.8°C) at 12 HPI (i.e. earliest measured timepoint post-infection), while P61 experienced the highest fever at any measured timepoint (41.3°C at 60 HPI) among all infected animals. They also had the highest lung histology scores among all transgenic pigs indicative of severe DAD (**Fig5.4 and 5.7**).

Typical symptoms of COVID-19 in humans involve fever, dry cough, fatigue and in severe cases, dyspnea. Most cases are however asymptomatic, especially in children and young adults. Elderly (>60 years) and people with co-morbidities are at a higher risk of severe disease, respiratory failure and death. The mean incubation period for SARS-CoV-2 is $\sim$ 5 days, with severe disease usually developing $\sim$ 8 days after symptom onset and critical disease or even death occurring at $\sim$ 16 days [333]. The SARS-CoV-2 infected transgenic pigs displayed similar mild-to-moderate clinical symptoms of COVID-19 in humans such as fever, coughing, sneezing and respiratory distress. The symptom onset in infected pigs is however comparatively early, detectable as soon as 1 DPI as shown in the second cohort and 2-3 DPI in the other cohorts. Infected animals also tested positive by lateral flow test as soon as 2 DPI (**Fig5.4**). Although some animals displayed more severe

symptoms, none resulted in fatal outcomes (although this was biased by the study design which aimed to avoid such outcomes). Infectious virus and viral RNA were detected at a relatively high level in the nasal turbinate of infected transgenic pigs at 2 and 4 DPI and dropped significantly at 7 DPI. This is consistent with the results reported in SARS-CoV-2 infected hamsters, where the level of both infectious virus and viral RNA in the nasal turbinate and lung also dropped significantly at 7 DPI [334]. Along with changes in the lungs suggestive of DAD, these virological, clinical and histological manifestations observed in pigs share some similarities with existing hamster models, which have a short incubation period, mild-to-moderate clinical symptoms despite moderate-to-severe lung pathology and signs of DAD in the lung. Unlike the K18-mouse model, where encephalitis is often the cause of death, we did not observe any neurological symptoms in the infected transgenic pigs.

Intriguingly, unlike mice, hamsters and NHP models, where a high level of SARS-CoV-2 RNA or even infectious virus can be recovered from the lungs[316, 334–336], a relatively lower level of viral RNA and no infectious virus was detected in the lungs of the infected transgenic pigs (**Fig 5.6**). This suggests the substantial inflammation and pathology observed in the lungs of infected pigs is driven by a dysfunctional host response, rather than damage caused by virus replication. This is consistent with the lungs of severe COVID-19 human patients, where the geographic distribution of viral RNA or protein was not linearly associated with pulmonary inflammatory changes within the lungs [331]. This further reflects the similarity in the anatomy, physiology and immune responses between humans and pigs, and this transgenic pig model may be more relevant than other existing models for understanding COVID-19-driven immunopathology. This would

be important for the evaluation of immunomodulatory and anti-inflammatory drugs.

Apart from pathogenesis, understanding the transmissibility of different SARS-CoV-2 variants is also crucial in evaluating the effect of vaccines in blocking transmission, a goal that is yet to be achieved. Although the hACE2-transgenic mouse model supports a certain degree of close contact transmission, the efficiency of respiratory droplet transmission was low [337]. While hamsters represent one of the best models for studying direct contact and aerosol transmission, transmission via fomite was not as efficient [338, 339]. In fact, the most popular model for studying the transmission of SARS-CoV-2 is ferrets, where viral spread is efficient through both direct and indirect contact. Nonetheless, they only present relatively mild disease following infection [314]. As high levels of infectious virus were recovered from the upper airways (i.e. nasal turbinate) and coughing and sneezing were observed, direct transmission between infected transgenic pigs would likely occur. Moreover, seroconversion was observed in at least one animal relatively early **(Fig5.8)**. Given the viral load, it is possible that all transgenic pigs would have seroconverted if the course of infection had run beyond seven days. The transgenic pig model could therefore serve as a transmission model for characterising vaccine efficacy, with animals exhibiting mild-to-moderate disease, recapitulating most COVID-19 cases in humans. Future experiments in testing indirect transmission among transgenic pigs would also be informative.

In humans, male sex and older age were associated with more severe disease caused by SARS-CoV-2 [340]. The transgenic pigs used in this challenge study are relatively young ($\sim$8 weeks old) and are mostly female (7/9 of infected animals), as the animals

selected were solely based on hACE2 expression (**Table5.1C**). Inference on demographic differences such as sex and age are not possible from this initial study. It will be interesting to test if infected transgenic pigs also display these differences as shown in the hamster models in the future, which can help to resolve the molecular mechanism behind these risk factors.

In this study, we have used three SARS-CoV-2 variants for the initial testing of the model. We acknowledge that these variants may not be clinically relevant anymore as the circulating variants are continuously evolving in the human population. As shown in **Fig5.3**, the susceptibility of the primary fibroblast cells towards early Omicron variant is substantially different from that of EDB2 and Delta, and it has been reported that later subvariants of Omicron have altered immunomodulatory potency. Therefore, it is likely that newer variants will behave differently in the same model. It will be important to test the susceptibility of the transgenic pigs towards newer variants *in vivo* in the future.

As discussed, adaptation has been reported when SARS-CoV-2 was passaged in non-adapted species both *in vitro* (e.g. Vero cells) and *in vivo* (e.g. mouse-adaptation). Organ-specific mutation has also been reported in humans [341]. Although not likely to have arisen after one passage, viral RNA isolated from the three tissue samples will be sequenced to determine whether mutations have occurred in the viral genome following infection of pigs.

CHAPTER 6

Conclusion and Future Direction

6.1 Conclusion

Zoonosis has long been a major threat to human health. However, most zoonotic events have only resulted in individual cases, with a small number of events managed to exacerbate and cause epidemics in specific areas. Although zoonotic coronaviruses have caused two smaller outbreaks in human in the past decade, it was not until the COVID-19 pandemic that the world truly came together and started appreciating their risk to public health. Before 2019, IAV had always been considered one of the viruses with the highest zoonotic potential, as demonstrated by numerous past epidemics and pandemics which caused substantial morbidity and mortality. Much research has been done to understand how this virus originated from birds managed to cross the species barrier and establish disease in a new host, which can help predict lineages that are more likely to cross the species barrier and cause disease in humans. Similarly, understanding the zoonotic barrier of coronavirus transmission can potentially help prevent a second COVID-19 pandemic.

The use of human ISG libraries for overexpression screens in the last decade has successfully identified a broad range of antiviral factors against adapted viruses [109, 288], while recent applications have identified factors specifically restricting non-adapted viruses, in particular IAV [181]. This suggests that IFN/ISG response plays a role in restricting cross-species transmission of IAV, and we hypothesise spill-over events of coronaviruses are also limited by such response.

The One Health concept suggests that human and animal health are interdependent,

as many zoonotic viruses were first transmitted to intermediate hosts that have close proximity to humans such as companion and livestock animals before causing disease in humans. Moreover, the transmission of viruses among different animal species can also cause substantial economic loss and undermine food security. Nonetheless, resources for high-throughput experiments to understand virus-host interaction in other animals are lacking. In light of this, our group has developed genome-wide CRISPR/Cas9 and lentiviral-based ISG libraries dedicated to pigs and chickens, two animal species that can act as hosts for potential zoonotic diseases with high economic value. In this project, we have applied the porcine ISG library to two coronaviruses, the pig-adapted PRCV and the non-adapted SARS-CoV-2, identifying novel porcine-specific restriction factors against coronaviruses.

At the beginning of the project, the aim was to develop detection methods for PRCV and characterised its replication in different porcine cell lines, as it is a relatively under-studied virus with limited available resources. We showed that ST and NPTr cells are highly susceptible to PRCV by an RT-qPCR-based assay. However, they are unsuitable for lentivirus transduction and cannot be used for the ISG screening experiment. In contrast, the less susceptible NSK cells are highly efficient for transduction. We therefore overexpressed the entry receptor for PRCV, pAPN, in NSK cells to generate pAPN-NSK cells which are highly susceptible to PRCV and highly transducible by lentiviral vectors. To enable screening against SARS-CoV-2, a similar approach was used to generate hACE2-NSK cells. We showed that an early SARS-CoV-2 isolate, EDB2 can infect hACE2-NSK cells more efficiently than the parental cell line, while the Omicron variant can only replicate in cells overexpressing the entry receptor. To quantify infectious viruses

for PRCV and EDB2, we optimised a TCID50 protocol using ST and hACE2-NSK cells respectively. Finally, we showed that PRCV and SARS-CoV-2 had different sensitivities to porcine IFN, specifically in NSK-derived cells, with the non-adapted virus being more sensitive.

This result led us to hypothesise that certain ISGs triggered by porcine IFN may contribute to barriers against SARS-CoV-2 infection of pigs *in vivo*. To identify general and cross-species porcine restriction factors against coronavirus, the porcine ISG library was applied to PRCV and SARS-CoV-2. Prior to the screening experiment, we optimised the use of an RT-qPCR-based direct lysis protocol as readout, as GFP-expressing virus and antibodies were not available for PRCV. The screens identified multiple porcine ISGs restricting both coronaviruses, with most also identified in previously published screens in human cells using the human library. These include ISGs involved in IFN signalling (i.e. IRF1, STAT2, cGAS, MDA5 etc.), the NF-κB pathway (i.e. TICAM2, RIPK1, RIPK2, MYD88 etc.) and antiviral factors with direct effect such as LGALS9 and OAS2, demonstrating the reliability of the screens. In addition, novel porcine-specific restriction factors against coronavirus, such as PHACTR1 and ZCWPW1, were also identified. We have validated these hits by TCID50, showing that they reduce infectious titres for both viruses. These data suggest although the IFN/ISG responses between humans and pigs are different, the factors responsible for upstream signalling are likely conserved.

The screens also identified porcine restriction factors that potentially restrict SARS-CoV-2 to a greater extent than PRCV, such as USF1, NRL and HOXD13. However, further experiments failed to validate this, with NRL not demonstrating a reduction in infectious

titres of either virus and USF1 and HOXD13 restricting both PRCV and SARS-CoV-2 potently. It was later found that this is likely due to differences in viral replication kinetics in the two cell lines, rather than intrinsic differences in inhibitory potency of USF1 against the two viruses. Nonetheless, USF1 and HOXD13 have not previously been demonstrated to have antiviral activities in published reports and are potential porcine-specific restriction factors.

Initial characterisation of ZCWPW1 coupled with available databases suggest it might be a tissue-specific ISG, as induction following IFN response is most robust in ST cells (a testis cell line) while being subtle and not-detected in NSK and porcine primary fibroblast cells respectively. However, inductions were measured 24 hours after treatment, with reports suggesting ZCWPW1 is induced much earlier. Attempts to test if the tissue-specificity of ZCWPW1 is conserved in other species have failed, as the mouse and human testis cell lines tested do not seem to respond to IFN, evidenced by the limited induction of marker ISG20. Surprisingly, siRNA knockdown of ZCWPW1 did not increase viral replication as expected for other antiviral factors. Rather, it resulted in a phenotype similar to overexpression, significantly reducing infectious PRCV titre. More importantly, this is not a result of cytotoxicity related to siRNA, as knowdown of ZCWPW1 did not affect cell viability. The fact that other groups had developed a ZCWPW1 knockout mouse model suggests this gene is not essential *in vivo* [294]. Preliminary data suggests overexpression of human ZCWPW1 in NSK-derived cells did not reduce PRCV or SARS-CoV-2 titre. However, due to time constraints, this experiment had only been performed once and repeats are needed to confirm this.

In addition to the IFN response, receptor compatibility and expressions remain a possible determinant of the lack of susceptibility of pigs to SARS-CoV-2 *in vivo*. To address this, we generated a transgenic pig model expressing hACE2. Prior to *in vivo* infection, primary fibroblast cells were isolated from selected animals and challenged with SARS-CoV-2 isolate EDB2, Delta and Omicron variants. The *in vitro* susceptibility to EDB2 and Delta correlates with hACE2 expression, with cells expressing more hACE2 producing more infectious virus and cells with undetectable levels not susceptible. A similar trend was observed in the case of Omicron, but only cells with the highest hACE2 expression are susceptible. This corresponds with the results shown in earlier chapters, where EDB2 can infect WT NSK cells while Omicron requires hACE2 overexpression. This suggests susceptibility to SARS-CoV-2, especially to the Omicron variant was not solely dependent on the expression of hACE2, and other barriers likely exist and contribute. As discussed, the BA.1 Omicron variant we use has less potency to evade innate immune suppression compared to earlier variants and later subvariants [260, 261, 332]. This possibly accounts for its atypical phenotype observed throughout the experiments in this project when compared to EDB2 or Delta.

As shown in **Fig 4.17A**, primary cells might have a stronger innate immune response when compared to cell lines such as NSK and ST cells as the porcine primary fibroblast cells elicit a stronger ISG20 induction following various stimulation. Taken together with the observation that hACE2 expression increases the susceptibility of cells to SARS-CoV-2, this could potentially explain why EDB2 can infect WT NSK cells (i.e. cell line with weaker innate response) without the need for hACE2 overexpression but could not infect primary cells unless hACE2 is present (i.e. stronger innate response of primary

cells overcame by the increased expression of entry receptor and therefore virus dosage). Meanwhile, Omicron can only infect hACE2-NSK cells (i.e. cell line with weaker innate immune response and increased susceptibility due to increased receptor expression) and only primary cells with the highest hACE2 expression (i.e. stronger innate response of primary cells overwhelmed by extensive entry receptor expression). This suggests that the porcine innate immunity may act as an alternate barrier, although it can be overcome by receptor overexpression.

Nine transgenic 8-week-old piglets expressing high-level hACE2 with three WT piglets were selected for the challenge study. Following SARS-CoV-2 infection, all transgenic animals showed symptoms consistent with mild-to-moderate COVID-19, including coughing, sneezing, loss of appetite and increased body temperature. More importantly, they all tested positive by LFT as soon as 2 DPI. WT pigs showed no symptoms and tested negative by LFT. High levels of SARS-CoV-2 RNA were detected in the nasal turbinate samples of all infected transgenic pigs and infectious virus were recovered with titre peaking at 4 DPI, but were undetectable by 7 DPI. In contrast, lower levels of viral RNA were detected in the trachea and lung samples with no infectious virus recovered.

Interestingly, the fibroblast cells isolated from P52 with a high hACE2 expression level were not susceptible to any SARS-CoV-2 variants infection *in vitro*. Similarly, although the animal itself was susceptible to EDB2 infection *in vivo*, displaying clinical symptoms and being LFT positive, the viral levels in all three tested tissues, along with the whole lung score of P52 were the lowest among all infected animals. Due to the scope of this project, this is not further investigated. However, it remains possible that some intrinsic

genetic features regarding the innate immune response in P52 might play a role in this.

Immunohistochemical staining confirmed the expression of hACE2 and SARS-CoV-2 nucleocapsid protein in the nasal turbinate, trachea and lung of infected transgenic pigs. Histological examination of tissues revealed clear signs of significant neutrophil and macrophage-rich inflammation within the lungs of infected animals from day four post-infection involving both bronchi and alveolar spaces with evidence of diffuse alveolar damage, oedema, and focal, fibrin-rich intravascular thrombi– all consistent with histological changes observed in fatal COVID-19.

The transgenic pig model has a similar physiology to humans and is highly susceptible to SARS-CoV-2, presenting mild-to-moderate symptoms with lung pathology consistent with severe COVID-19 patients. This large animal model can act as an alternative to existing rodents and NHP models to help develop improved vaccines and therapeutics against SARS-CoV-2.

6.2 Future Direction

As discussed in Chapter 3, one of the major aims of this project is to compare the effect of porcine and human ISGs against coronaviruses. However, our attempt to generate human cell lines susceptible to PRCV infection for comparative screening was unsuccessful. Apart from repeating the attempt, the use of another porcine coronavirus is an option. SADS-CoV has relatively recently been identified to cause disease in pigs *in vivo* while also having the ability to infect humans and a wide range of cells from other species *in vitro* [219, 342], complementary to SARS-CoV-2 which causes disease in humans *in vivo* and infects pig cells *in vitro*. With the optimised direct lysis RT-qPCR protocol for read-out, which only requires primer pair design, screens against SADS-CoV in human and pig cells using the human and porcine ISG libraries can be performed with relative ease, which can potentially identify human-specific ISG restricting non-adapted coronavirus. The minimal optimisation required by the direct lysis protocol also allowed screens against undiscovered or other less-studied viruses where antibodies are not readily available.

Several novel hits had been identified and validated from the porcine ISG screens against PRCV and SARS-CoV-2. However, due to time constraints, it is not possible to characterise them all in this project. Nonetheless, the initial characterisation for ZCWPW1 set a solid foundation for further investigation. We have detected an upregulation of ZCWPW1 following IFN stimulation at the mRNA level. However, we cannot access it at the protein level as the commercially available antibodies targeting human and murine ZCWPW1 do not recognise porcine ZCWPW1 by western blot. Generating antibodies targeting porcine ZCWPW1 or tagging the gene by CRISPR/Cas9 knockin is required

to study its endogenous regulation at the protein level. Meanwhile, HA-tagged human and porcine ZCWPW1 constructs had been produced. This allows future experiments to study the localisation of ZCWPW1 and potential interaction with other host or viral proteins.

The knockdown of ZCWPW1 by siRNA resulted in a phenotype similar to overexpression, reducing PRCV replication. To confirm this phenotype, ZCWPW1 can be knocked out by CRISPR/Cas9. Published reports on MORC3 provided insights into the mechanism of action for the phenotype of ZCWPW1. This antiviral ISG which also possesses the zinc finger CW-type domain triggers IFN/ISG responses when degraded. Expression of IFN and other marker ISG can be measured following knockdown or knockout of ZCWPW1 to investigate if this is also the mechanism of action for porcine ZCWPW1. Preliminary data suggests human ZCWPW1 had little to no antiviral properties in porcine cells. However, this may be due to the incompatibility of human factors in a porcine background. To confirm the species-specific phenotype, porcine and human ZCWPW1 will be expressed in a human cell line and challenged with SARS-CoV-2. If the observed phenotype remains unchanged, differences in amino acid sequence between human and porcine ZCWPW1 will most probably account for this. Chimeric ZCWPW1 can then be generated to investigate which region in porcine ZCWPW1 is responsible for its antiviral activity.

The goal of the pilot *in vivo* challenge study in this project is to test if the generated pigs are susceptible to SARS-CoV-2, and we acknowledge that some other important aspects for characterising this model were not taken into account. For example, innate,

humoral and cellular immune responses, as well as other clinical measurements such as oxygen saturation levels and respiratory rate can be taken in future studies. Later time points should also be included as peak viral replication in the lung and severe disease processes are thought to progress more significantly in the second week of infection.

The early EDB2 isolate was chosen for the initial study as it efficiently infected primary fibroblast cells from the transgenic pigs and early isolates were associated with more severe disease, therefore more likely to generate clear clinical and histological outcomes in the challenge study. However, it is not clinically relevant anymore as newer variants and subvariants have replaced it. Although we have tested the susceptibility of the primary fibroblast cells to more recent variants *in vitro*, a new challenge study is needed to confirm their *in vivo* susceptibility, especially for Omicron, before the model can be applied to SARS-CoV-2 research related to vaccines and therapeutic development.

Pigs are not naturally susceptible and we showed that overexpression of hACE2 renders them susceptible to SARS-CoV-2 *in vivo*. This suggests expression of receptors plays a key role in determining susceptibility. However, it is not likely due to receptor compatibility, as several porcine cell lines were susceptible and pACE2 can bind SARS-CoV-2 spike proteins. While enhanced apoptosis in porcine primary respiratory epithelial cells in contrast to human cells was hypothesised to contribute to self-limiting SARS-CoV-2 replication in pigs [343], it is likely that other barriers exist and contribute to a lack of susceptibility.

As discussed earlier, unlike EDB2 and the Delta variant, the susceptibility of the primary fibroblast cells from transgenic animals to the Omicron variant was not solely dependent

on the expression of hACE2, as all except the two cell lines with the highest expression level show susceptibility. The weaker innate immune suppression of BA.1 used in this project possibly accounts for its phenotype and required an optimal condition (i.e. weaker innate response coupled with increased entry receptor expression) to infect pig cells. Infecting the transgenic primary fibroblast cells with more recent Omicron subvariants, which have evolved to better suppress the human innate response, could help address this hypothesis.

The unusual phenotype observed in P52 should also be followed up as resolving the molecular mechanism of resistance of pigs to SARS-CoV-2 infection may hold therapeutic value for the treatment of COVID-19. Primary fibroblast cells isolated from this animal are not susceptible to any SARS-CoV-2 variants despite substantial hACE2 expression, which correlates with the lowest viral level detected in its tissue among all other infected animals. The hACE2 insert should first be sequenced to confirm such lack of susceptibility is not due to spontaneous mutation in the entry receptor. Mechanistic studies can also be carried out in the isolated cells to identify at which stage the virus gets restricted (e.g. entry, post-entry, transcription and translation or egress). This may help narrow the potential genetic targets for comparative analysis of the genomes of P52 and other transgenic animals, and potentially help in identifying the underlying genetic determinants in P52 contributing to its unusual disease-susceptibility phenotype.

One major challenge in studying pathogen-host interaction in farm animal species or other non-model organisms is the relative lack of literature and resources. Although sequencing techniques and subsequently the quality of genomic data have improved

substantially in the past decades, the main limitation remains in the annotation of these genomes. Many genes in the genome of non-model organisms have been incorrectly annotated, affecting our interpretation of their potential functions. For example, recent phylogenetic analyses suggest the currently annotated chicken IRF3 are more closely related to mammalian IRF7, suggesting chickens likely lack mammalian IRF3 homologue [344, 345]. Similarly, the currently annotated chicken IRF9 is both genetically and phylogenomically more similar to IRF10, with speculation that IRF9 is also absent in chickens [346]. As discussed in the first chapter, IRF3 is constitutively expressed in humans and is responsible for early IFN induction; while IRF9 is a component of ISGF3 involved in ISG induction. These potential misannotations raised concerns, highlighting a fundamental difference between the innate immune systems of birds and humans. This is further exemplified by the fact that our group have experienced substantially more challenges when working with the chicken CRISPR and ISG libraries than the porcine libraries. More research into the genome of farm animals is anticipated with increased concerns over them as potential hosts for zoonotic diseases.

Appendix

Complete list of genes included in the porcine ISG library

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
1	PARP11	ENSSSCG00000000728	1102
2	TBX3	ENSSSCG00000009865	2284
3	TNFSF13B	ENSSSCG00000009542	748
4	MBOAT2	ENSSSCG00000023539	1609
5	PANX1	ENSSSCG00000032360	1324
6	KLHL5	ENSSSCG00000008787	1771
7	NMI	ENSSSCG00000022011	838
8	STX11	ENSSSCG00000004125	910
9	SULT1C4	ENSSSCG00000025836	979
10	PCDH19	ENSSSCG00000012479	3484
11	ISG20	ENSSSCG00000024867	562
12	PSMB10	ENSSSCG00000031781	937
13	IL4I1	ENSSSCG00000003192	1780
14	K7GNK3	ENSSSCG00000024161	1117
15	TOR1B	ENSSSCG00000005683	1036
16	NUP62CL	ENSSSCG00000012557	967
17	JCAD	ENSSSCG00000011020	4015
18	BLNK	ENSSSCG00000010504	1438
19	TMEM100	ENSSSCG00000017607	451
20	A0A287AHT6	ENSSSCG00000003311	730
21	TCP11L2	ENSSSCG00000024881	1606
22	N/A	ENSSSCG00000001231	769
23	RNF103	ENSSSCG00000008218	2104
24	ZCCHC2	ENSSSCG00000004897	3580
25	D8QQJ2	ENSSSCG00000021236	787
26	MLKL	ENSSSCG00000002716	1384
27	ZNF618	ENSSSCG00000005484	2803
28	SYN1	ENSSSCG00000012276	2185
29	IDO1	ENSSSCG00000007007	1336
30	ATAD2	ENSSSCG00000005983	4201
31	MAD2L2	ENSSSCG00000003419	682
32	F1RTI3	ENSSSCG00000001394	1063
33	STAT1	ENSSSCG00000016057	2416
34	SOCS3-like	ENSSSCG00000038659	655
35	STK17A	ENSSSCG00000016758	1285
36	USP15	ENSSSCG00000000457	2992
37	TRIM22	ENSSSCG00000014670	1525
38	GM2A	ENSSSCG00000017087	628

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
39	IKZF2	ENSSSCG00000016164	1645
40	PPP1R1B	ENSSSCG00000017498	661
41	PLAG1	ENSSSCG00000006247	1546
42	LRRN3	ENSSSCG00000016652	2173
43	PRRG1	ENSSSCG00000027777	703
44	SAT1-like	ENSSSCG00000002541	484
45	ALPK1-like	ENSSSCG00000039387	2350
46	ACSL1	ENSSSCG00000015784	2146
47	XRN1	ENSSSCG00000011678	793
48	RNASEL	ENSSSCG00000015549	2278
49	IFITM3	ENSSSCG00000038912	484
50	RNF19A	ENSSSCG00000006066	2563
51	OTUD4	ENSSSCG00000009042	3379
52	RAB27A	ENSSSCG00000004612	712
53	ANKRD37	ENSSSCG00000015798	565
54	TMEM173	ENSSSCG00000014349	1300
55	RASGEF1B	ENSSSCG00000029507	1726
56	IFITM1	ENSSSCG00000014565	430
57	MT2A	ENSSSCG00000030300	232
58	PLAC8	ENSSSCG00000009240	388
59	SEMA7A	ENSSSCG00000023165	2197
60	ARL4A	ENSSSCG00000032709	649
61	PDE12	ENSSSCG00000011465	1876
62	cGAS	ENSSSCG00000021383	1534
63	ADAR	ENSSSCG00000006543	3463
64	DEPP1	ENSSSCG00000022256	673
65	PCGF5	ENSSSCG00000036083	817
66	AFAP1L1	ENSSSCG00000014431	2335
67	STEAP4	ENSSSCG00000015299	1459
68	TM4SF1	ENSSSCG00000033018	655
69	MAB21L2	ENSSSCG00000009022	1126
70	HERC6	ENSSSCG00000021712	3112
71	TAL1	ENSSSCG00000035847	1042
72	MXD1	ENSSSCG00000008334	712
73	TICAM2	ENSSSCG00000038141	760
74	PAX2	ENSSSCG00000010556	1309
75	BAK1	ENSSSCG00000001516	682
76	GBP1	ENSSSCG00000024973	1822
77	PML	ENSSSCG00000001912	2668
78	MX1	ENSSSCG00000012077	2032
79	DLG3	ENSSSCG00000012375	2596

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
80	CAMK2D	ENSSSCG00000009123	1384
81	TLR6	ENSSSCG00000026592	2437
82	WARS1	ENSSSCG00000002516	1495
83	TDRD7	ENSSSCG00000005364	3343
84	SEMA4D	ENSSSCG00000009584	2644
85	TCF7L2	ENSSSCG00000010638	2005
86	DDX58 (RIG-I)	ENSSSCG00000030408	2878
87	TEC	ENSSSCG00000008820	1945
88	IL15RA	ENSSSCG00000021815	571
89	TSPAN13	ENSSSCG00000015362	661
90	ERAP1	ENSSSCG00000014171	2872
91	TRIM38	ENSSSCG00000035153	1444
92	PELP1	ENSSSCG00000017924	3373
93	TMEM106A	ENSSSCG00000025856	832
94	TMEM150C	ENSSSCG00000026729	877
95	ENTPD4	ENSSSCG00000031871	976
96	TGM1	ENSSSCG00000001993	2527
97	RHBDL3	ENSSSCG00000017734	1372
98	IFI6	ENSSSCG00000034570	448
99	APOBEC3B	ENSSSCG00000000091	1411
100	CCDC126	ENSSSCG00000029174	469
101	SDCBP	ENSSSCG00000006237	1063
102	CNIH4	ENSSSCG00000035034	466
103	RALGDS	ENSSSCG00000005738	2701
104	PLEKHA4	ENSSSCG00000003137	2131
105	IFI35	ENSSSCG00000029763	901
106	CXCL12	ENSSSCG00000034973	397
107	CCRL2	ENSSSCG00000023557	1099
108	FAM171B	ENSSSCG00000025134	2512
109	B2M	ENSSSCG00000004687	469
110	IRF1	ENSSSCG00000014277	1015
111	SNX25	ENSSSCG00000015789	2185
112	FZD6	ENSSSCG00000006052	2182
113	ZCWPW1	ENSSSCG00000007659	1990
114	CLEC14A	ENSSSCG00000023760	1528
115	N4BP1	ENSSSCG00000002841	2740
116	EBF1	ENSSSCG00000017046	1822
117	EBF3	ENSSSCG00000010757	2557
118	BATF3	ENSSSCG00000026587	415
119	TRIM6	ENSSSCG00000032133	1516
120	SLCO5A1	ENSSSCG00000006196	1411

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
121	GBP2	ENSSSCG00000006923	1819
122	TRAF3IP2	ENSSSCG00000004420	1744
123	TTC31	ENSSSCG00000008275	1606
124	NCF1	ENSSSCG00000007722	1225
125	GBP6-like	ENSSSCG00000035732	625
126	TAPBP	ENSSSCG00000001507	1567
127	ATP10A	ENSSSCG00000004830	4498
128	MORC3	ENSSSCG00000012055	2503
129	SLC2A6	ENSSSCG00000024166	1567
130	UNC93B1	ENSSSCG00000012893	2197
131	RASSF4	ENSSSCG00000029662	1048
132	NRL	ENSSSCG00000002010	835
133	TENT4B	ENSSSCG00000040887	2179
134	CASP1	ENSSSCG00000028711	1267
135	C19orf66	ENSSSCG00000013664	904
136	PNKD	ENSSSCG00000029002	1558
137	VCAM1	ENSSSCG00000006862	1660
138	RND3	ENSSSCG00000040445	781
139	EDNRA	ENSSSCG00000009031	1330
140	FAM76A	ENSSSCG00000027894	967
141	NDC80	ENSSSCG00000003697	1975
142	RNF19B	ENSSSCG00000025206	2242
143	INTS2	ENSSSCG00000017673	3637
144	LGALS8	ENSSSCG00000010146	1117
145	SNX10	ENSSSCG00000037598	649
146	IFIT5	ENSSSCG00000010454	1495
147	CDC14A	ENSSSCG00000006864	1885
148	VASH2	ENSSSCG00000015589	1114
149	NKX3-1	ENSSSCG00000031924	790
150	MRAP2	ENSSSCG00000026943	682
151	RNF152	ENSSSCG00000004902	658
152	PDE12-like	ENSSSCG00000034476	1321
153	ENPP4	ENSSSCG00000025788	1408
154	CTSS	ENSSSCG00000006648	1042
155	SAMD9	ENSSSCG00000027372	4837
156	ZFP36L1	ENSSSCG00000028304	1063
157	NBN	ENSSSCG00000006127	2305
158	ARHGEF3	ENSSSCG00000030359	1756
159	SP140-like	ENSSSCG00000016263	1375
160	MSX1	ENSSSCG00000008727	958
161	MASTL	ENSSSCG00000011065	2674

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
162	STN1	ENSSSCG00000010605	1360
163	F1SV36	ENSSSCG00000014997	1063
164	ATF3	ENSSSCG00000015595	799
165	DMP1	ENSSSCG00000009220	1579
166	CSF1	ENSSSCG00000037449	1711
167	MAX	ENSSSCG00000037267	529
168	NEMP2	ENSSSCG00000016053	430
169	EPSTI1	ENSSSCG00000037572	961
170	PSMB9	ENSSSCG00000001463	706
171	VWA5A	ENSSSCG00000015175	2521
172	DIO2	ENSSSCG00000040638	856
173	TRIM14	ENSSSCG00000033222	1384
174	SAT1	ENSSSCG00000012173	748
175	SAMHD1	ENSSSCG00000027806	1930
176	ARL6IP1	ENSSSCG00000025281	658
177	RIPK1	ENSSSCG00000001009	2041
178	OASL	ENSSSCG00000009921	1093
179	CMTR1	ENSSSCG00000001579	2236
180	SFMBT2	ENSSSCG00000024480	2746
181	IL10RB	ENSSSCG00000036556	1765
182	RHPN2	ENSSSCG00000022312	1654
183	EDN1	ENSSSCG00000035598	658
184	ASPHD2	ENSSSCG00000009959	1135
185	ETV6	ENSSSCG00000024152	1405
186	STAT2	ENSSSCG00000000396	2641
187	TREX1	ENSSSCG00000011358	991
188	ZNF791-LIKE	ENSSSCG00000029347	1888
189	CBR3	ENSSSCG00000012053	880
190	IFITM1-LIKE	ENSSSCG00000032591	466
191	GBP-6	ENSSSCG00000030801	2302
192	FBXO39	ENSSSCG00000017886	1894
193	ITSN1	ENSSSCG00000024614	4513
194	ATAD1	ENSSSCG00000010438	1132
195	SERPING1	ENSSSCG00000013181	1522
196	PODXL	ENSSSCG00000016548	1378
197	CTSO	ENSSSCG00000008867	982
198	RAB8B	ENSSSCG00000026571	670
199	ERAP2	ENSSSCG00000027860	2899
200	CBLN3	ENSSSCG00000001985	664
201	MR1	ENSSSCG00000015540	1060
202	ARID5A	ENSSSCG00000008123	1840

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
203	SLA-5	ENSSSCG00000030767	1123
204	SIRT1	ENSSSCG00000010235	2275
205	JAK2	ENSSSCG00000005215	3442
206	MGAT3	ENSSSCG00000038126	1648
207	AP1S3	ENSSSCG00000040847	511
208	RASGRP3	ENSSSCG00000008509	2125
209	MAFF	ENSSSCG00000030165	595
210	SETDB2	ENSSSCG00000009395	2149
211	TNFSF10 (TRAIL)	ENSSSCG00000020906	955
212	ARHGAP42	ENSSSCG00000014976	2671
213	C5orf30	ENSSSCG00000038779	670
214	LACTB	ENSSSCG00000004569	1798
215	NCOA1	ENSSSCG00000008581	4372
216	SERPINB1	ENSSSCG00000001011	1288
217	CXCL10	ENSSSCG00000032474	361
218	ICAM1	ENSSSCG00000013655	1660
219	TEAD4	ENSSSCG00000000736	1429
220	ELF1	ENSSSCG00000009440	1897
221	MUC1	ENSSSCG00000006525	1963
222	ABCC1-LIKE	ENSSSCG00000012006	3937
223	CEP44	ENSSSCG00000032475	1204
224	PARP12	ENSSSCG00000016502	2155
225	CPNE8	ENSSSCG00000000778	1780
226	C1QTNF3	ENSSSCG00000016823	1006
227	IPO13	ENSSSCG00000029430	2791
228	CBX4	ENSSSCG00000017160	1690
229	IFIT1	ENSSSCG00000010452	1576
230	SMAGP	ENSSSCG00000036904	379
231	TOR3A	ENSSSCG00000015525	1261
232	PXK	ENSSSCG00000011474	1780
233	MIS18BP1	ENSSSCG00000004999	3226
234	SLA-11	ENSSSCG00000001341	1156
235	MYD88	ENSSSCG00000011251	994
236	TNFAIP3	ENSSSCG00000004154	2404
237	TLR2	ENSSSCG00000009002	2404
238	PIK3R3	ENSSSCG00000003909	1432
239	IRF8	ENSSSCG00000034980	1327
240	KCNJ8	ENSSSCG00000000573	1321
241	NUAK2	ENSSSCG00000038471	1834
242	CNDP2	ENSSSCG00000004870	1474
243	SAP30L	ENSSSCG00000033176	595

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
244	NKX2-2	ENSSSCG00000040017	868
245	ZBP1	ENSSSCG00000007508	1363
246	TIFA	ENSSSCG00000009129	655
247	PSME2-like	ENSSSCG00000017618	604
248	DDX60	ENSSSCG00000009720	5161
249	PSME2	ENSSSCG00000002004	766
250	TAP1	ENSSSCG00000025618	2287
251	SLC22A15	ENSSSCG00000006743	1696
252	UBA7	ENSSSCG00000011393	3085
253	ALPK1	ENSSSCG00000031348	475
254	RIPOR3	ENSSSCG00000007470	2890
255	NAMPT	ENSSSCG00000015435	1621
256	PHACTR4	ENSSSCG00000028973	2116
257	ISG15	ENSSSCG00000040575	550
258	CASP8	ENSSSCG00000028157	1507
259	DTX3L	ENSSSCG00000011876	2152
260	ISG12(A)	ENSSSCG00000035297	643
261	FAM111B	ENSSSCG00000013147	2245
262	TTF2	ENSSSCG00000006733	3547
263	PSMB8	ENSSSCG00000026951	934
264	IRF9	ENSSSCG00000002002	1282
265	STARD8	ENSSSCG00000012828	3388
266	C3orf14	ENSSSCG00000011486	481
267	FAM111A	ENSSSCG00000033703	1864
268	DAXX	ENSSSCG00000001509	2320
269	NXPE3	ENSSSCG00000011950	1726
270	TBC1D10A (EPI64)	ENSSSCG00000009999	1684
271	CD274	ENSSSCG00000005211	910
272	TNFAIP8	ENSSSCG00000040617	685
273	GBP7	ENSSSCG00000006919	1399
274	GUCY1A1	ENSSSCG00000008866	2113
275	TRIM56	ENSSSCG00000034347	2314
276	VWCE	ENSSSCG00000013093	2872
277	GPR63	ENSSSCG00000035240	1306
278	USF1	ENSSSCG00000006371	979
279	MNDA	ENSSSCG00000006418	1804
280	MOV10	ENSSSCG00000006776	3298
281	CLEC2D	ENSSSCG00000000657	685
282	CCL5	ENSSSCG00000017705	322
283	TIGAR	ENSSSCG00000024219	856
284	TMEM243	ENSSSCG00000025326	403

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
285	ETV7	ENSSSCG00000001561	1051
286	REC8	ENSSSCG00000002001	1789
287	TBPL1	ENSSSCG00000022387	607
288	NUB1	ENSSSCG00000016438	1966
289	TLR3	ENSSSCG00000015801	2755
290	MAK	ENSSSCG00000001042	2125
291	TRAFD1	ENSSSCG00000009886	1792
292	SPSB4	ENSSSCG00000034369	820
293	CD38	ENSSSCG00000008742	907
294	FUT8	ENSSSCG00000002283	1792
295	VCPIP1	ENSSSCG00000006209	3718
296	PPP1R16B	ENSSSCG00000007350	1753
297	HES4	ENSSSCG00000035420	733
298	CCL11	ENSSSCG00000040648	346
299	KCNA1	ENSSSCG00000000716	1534
300	RFTN2	ENSSSCG00000037483	1570
301	CCL2	ENSSSCG00000017723	346
302	C2	ENSSSCG00000001422	2608
303	FMR1	ENSSSCG00000012727	1945
304	SP140	ENSSSCG00000016262	2305
305	HMCES	ENSSSCG00000011611	1150
306	BATF2	ENSSSCG00000023178	877
307	EIF2AK2 (PKR)	ENSSSCG00000008496	1660
308	ZFPM2	ENSSSCG00000006038	3502
309	IFI44	ENSSSCG00000003763	1363
310	NGLY1	ENSSSCG00000011214	2011
311	ARL5B	ENSSSCG00000031456	589
312	CHST2 (GST2)	ENSSSCG00000022031	1441
313	IFI44L	ENSSSCG00000027660	1303
314	PLEKHH2	ENSSSCG00000008456	4708
315	SLC25A28	ENSSSCG00000037536	1141
316	MHC1-126-LIKE	ENSSSCG00000001229	1231
317	BST2	ENSSSCG00000033453	580
318	ZNF710	ENSSSCG00000021361	2467
319	MAP3K14	ENSSSCG00000017330	2998
320	HRH1	ENSSSCG00000011576	1543
321	HERC5	ENSSSCG00000030548	2953
322	MAN1A1	ENSSSCG00000004244	2026
323	CIITA	ENSSSCG00000007901	3544
324	UBE2L6	ENSSSCG00000023379	508
325	PARP9	ENSSSCG00000027709	2590

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
326	TAGAP	ENSSSCG00000004053	2563
327	RIPK2	ENSSSCG00000024096	1648
328	WASHC4	ENSSSCG00000000840	3565
329	EPS8	ENSSSCG00000000600	2554
330	MX2	ENSSSCG00000012076	2182
331	COBLL1	ENSSSCG00000025598	4222
332	CD40	ENSSSCG00000035078	883
333	CRY1	ENSSSCG00000000164	1813
334	FUNDC1	ENSSSCG00000012258	523
335	TRIM25	ENSSSCG00000017614	1945
336	SOCS1	ENSSSCG00000027855	709
337	TAPBPL	ENSSSCG00000000704	1372
338	TRIM26	ENSSSCG00000001233	1681
339	CLIC2	ENSSSCG00000039998	790
340	SOCS3	ENSSSCG00000036956	736
341	TLK2	ENSSSCG00000017301	2299
342	TRIM21	ENSSSCG00000014780	1453
343	RUBCN	ENSSSCG00000011853	2962
344	IL7	ENSSSCG00000006161	577
345	MFSD14B	ENSSSCG00000026552	1645
346	DCP2	ENSSSCG00000014204	1318
347	S100A14	ENSSSCG00000006582	361
348	VSTM5	ENSSSCG00000014950	757
349	SLC15A3	ENSSSCG00000013114	1768
350	PRRG4	ENSSSCG00000013313	727
351	GDAP2	ENSSSCG00000006728	1270
352	RNASEH2B	ENSSSCG00000024385	886
353	TENT5A	ENSSSCG00000004464	1351
354	TBC1D32	ENSSSCG00000004242	3949
355	IFITM1-like	ENSSSCG00000032436	424
356	PI4K2B	ENSSSCG00000027340	1843
357	IFIT2	ENSSSCG00000010451	1468
358	ZNF709	ENSSSCG00000013717	1717
359	AKAP7	ENSSSCG00000036790	1330
360	USP6NL	ENSSSCG00000011120	2506
361	CASP10	ENSSSCG00000026940	1609
362	C1S	ENSSSCG00000039847	2128
363	SLC2A12	ENSSSCG00000004172	1912
364	ANKFY1	ENSSSCG00000017877	3559
365	FLT3LG	ENSSSCG00000003167	754
366	LMO2	ENSSSCG00000013307	718

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
367	PPA1	ENSSSCG00000010261	916
368	TRAF1	ENSSSCG00000005511	1294
369	PRRG1-like	ENSSSCG00000036838	343
370	HOXD13	ENSSSCG00000015979	1072
371	STOML1	ENSSSCG00000001913	1411
372	DHX58 (LGP2)	ENSSSCG00000017416	2092
373	3LGG5	ENSSSCG00000014672	1258
374	VPS54	ENSSSCG00000027509	2980
375	TMEM140	ENSSSCG00000040673	607
376	WDR76	ENSSSCG00000004695	2083
377	MT1A	ENSSSCG00000023684	232
378	CALN1	ENSSSCG00000007733	2449
379	BCL2L14	ENSSSCG00000000623	1036
380	TCIRG1	ENSSSCG00000012890	2548
381	RHBDF2	ENSSSCG00000023362	2530
382	CDHR4	ENSSSCG00000011391	2383
383	ANGPTL1	ENSSSCG00000015522	1555
384	PMAIP1	ENSSSCG00000026454	412
385	TCF21	ENSSSCG00000033258	586
386	OAS2	ENSSSCG00000009881	2098
387	BID	ENSSSCG00000000769	664
388	GNG11	ENSSSCG00000015324	268
389	CMPK2	ENSSSCG00000008647	1396
390	RGS2	ENSSSCG00000037241	685
391	GDAP1	ENSSSCG00000006173	1135
392	PCGF5-like	ENSSSCG00000040588	760
393	MED25	ENSSSCG00000003197	1669
394	CMAH	ENSSSCG00000001099	1780
395	RFX5	ENSSSCG00000006625	2053
396	TMTC1	ENSSSCG00000027447	2920
397	GATA6	ENSSSCG00000003702	1825
398	SCO2	ENSSSCG00000031420	841
399	SLC37A2	ENSSSCG00000022536	1555
400	TBC1D22B	ENSSSCG00000001577	1738
401	RSAD2 (Viperin)	ENSSSCG00000008648	1135
402	TAP2	ENSSSCG00000025593	2314
403	PHACTR1	ENSSSCG00000001052	1816
404	SERPINE1	ENSSSCG00000025698	1252
405	TSC22D1	ENSSSCG00000009422	3259
406	HEG1	ENSSSCG00000011859	4009
407	GLUL	ENSSSCG00000015545	1192

Identifier (SSC)	Gene Name	Ensmebl ID	Size (bp)
408	APOL3-like	ENSSSCG00000000148	1066
409	LGALS9	ENSSSCG00000017754	1114
410	CNP	ENSSSCG00000017420	1309
411	OSMR	ENSSSCG00000016851	2980
412	ASB9	ENSSSCG00000012132	838
413	PLEKHA7	ENSSSCG00000013382	3940
414	ZC3HAV1 (ZAP)	ENSSSCG00000016512	2731
415	PTGIR	ENSSSCG00000026602	1201
416	IFIH1 (MDA5)	ENSSSCG00000015897	3118
417	LPAR6	ENSSSCG00000036063	1084
418	USP18	ENSSSCG00000000774	1054
419	JADE2	ENSSSCG00000014303	2596
420	DLL4	ENSSSCG00000004755	2104
421	GMPR	ENSSSCG00000001064	1099
422	LAP3	ENSSSCG00000023604	1606
423	INSIG1	ENSSSCG00000016420	973
424	DUSP6	ENSSSCG00000037910	1192
425	IP6K1	ENSSSCG00000011390	1372
426	SLC12A7	ENSSSCG00000034302	3511
427	TBX18	ENSSSCG00000004290	1885
428	ATP10D	ENSSSCG00000008812	4318
429	IRF7	ENSSSCG00000012853	1510
430	CREM	ENSSSCG00000011106	904
431	IRF2	ENSSSCG00000015782	1324
432	SP110	ENSSSCG00000016261	2197

Bibliography

1. Baltimore, D. Expression of animal virus genomes. *Bacteriological Reviews* **35,** 235–241 (1971).
2. Galibert, F., Mandart, E., Fitoussi, F., Tiollais, P. & Charnay, P. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli. *Nature* **281,** 646–650 (1979).
3. Gorbalenya, A. E. *et al.* The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks. *Nature Microbiology* **5,** 668–674 (2020).
4. Bukhari, K. *et al.* Description and initial characterization of metatranscriptomic nidovirus-like genomes from the proposed new family Abyssoviridae, and from a sister group to the Coronavirinae, the proposed genus Alphaletovirus. *Virology* **524,** 160–171 (2018).
5. Mordecai, G. J. *et al.* Endangered wild salmon infected by newly discovered viruses. *eLife* **8** (2019).
6. Woo, P. C. Y. *et al.* ICTV Virus Taxonomy Profile: Coronaviridae 2023. *Journal of General Virology* **104** (2023).
7. Woo, P. C. Y. *et al.* Discovery of Seven Novel Mammalian and Avian Coronaviruses in the Genus Deltacoronavirus Supports Bat Coronaviruses as the Gene Source of Alphacoronavirus and Betacoronavirus and Avian Coronaviruses as the Gene Source of Gammacoronavirus and Deltacoronavirus. *Journal of Virology* **86** (2012).
8. Schalk, A. & Hawn, M. An apparently new respiratory disease of baby chicks. *Journal of the American Veterinary Medical Association,* 413–423 (1931).
9. DOYLE, L. P. & HUTCHINGS, L. M. A transmissible gastroenteritis in pigs. *Journal of the American Veterinary Medical Association* **108,** 257–9 (1946).
10. Gledhill, A. & Andrewes, C. A hepatitis virus of mice. *British journal of experimental pathology* **32,** 559–568 (1951).

11. Almeida, J. D. & Tyrrell, D. A. J. The Morphology of Three Previously Uncharacterized Human Respiratory Viruses that Grow in Organ Culture. *Journal of General Virology* **1,** 175–178 (1967).
12. Almeida, J. *et al.* Virology: Coronaviruses. *Nature* **220,** 650–650 (1968).
13. Brinton, M. A. & Miller, W. A. Positive strand RNA virus replication: It depends on the ends. *Virus Research* **206,** 1–2 (2015).
14. Knipe, D. M. & Howley, P. *Fields virology: Sixth edition* (2013).
15. De Haan, C. A. M., Volders, H., Koetzner, C. A., Masters, P. S. & Rottier, P. J. M. Coronaviruses Maintain Viability despite Dramatic Rearrangements of the Strictly Conserved Genome Organization. *Journal of Virology* **76,** 12491–12502 (2002).
16. Woo, P. C., Lau, S. K., Huang, Y. & Yuen, K. Y. *Coronavirus diversity, phylogeny and interspecies jumping* 2009.
17. Liu, D. X., Fung, T. S., Chong, K. K.-L., Shukla, A. & Hilgenfeld, R. Accessory proteins of SARS-CoV and other coronaviruses. *Antiviral Research* **109,** 97–109 (2014).
18. Zeng, Q., Langereis, M. A., van Vliet, A. L. W., Huizinga, E. G. & de Groot, R. J. Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proceedings of the National Academy of Sciences* **105,** 9065–9069 (2008).
19. Langereis, M. A., Zeng, Q., Heesters, B., Huizinga, E. G. & de Groot, R. J. The Murine Coronavirus Hemagglutinin-esterase Receptor-binding Site: A Major Shift in Ligand Specificity through Modest Changes in Architecture. *PLoS Pathogens* **8,** e1002492 (2012).
20. Zhang, X. *et al.* Complete genomic sequences, a key residue in the spike protein and deletions in nonstructural protein 3b of US strains of the virulent and attenuated coronaviruses, transmissible gastroenteritis virus and porcine respiratory coronavirus. *Virology* (2007).
21. De Haan, C. A., Masters, P. S., Shen, X., Weiss, S. & Rottier, P. J. The Group-Specific Murine Coronavirus Genes Are Not Essential, but Their Deletion, by Reverse Genetics, Is Attenuating in the Natural Host. *Virology* **296,** 177–189 (2002).
22. Gorbalenya, A. E. in *Nidoviruses* (2014).
23. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology* **19** (2021).
24. Curtis, K. M., Yount, B., Sims, A. C. & Baric, R. S. Reverse Genetic Analysis of the Transcription Regulatory Sequence of the Coronavirus Transmissible Gastroenteritis Virus. *Journal of Virology* **78,** 6061–6066 (2004).
25. Tortorici, M. A. & Veesler, D. in, 93–116 (2019).

26. Irigoyen, N. *et al.* High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. *PLOS Pathogens* **12,** e1005473 (2016).
27. Finkel, Y. *et al.* The coding capacity of SARS-CoV-2. *Nature* **589,** 125–130 (2021).
28. Brierley, I., Digard, P. & Inglis, S. C. Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. *Cell* (1989).
29. Perlman, S. & Netland, J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nature Reviews Microbiology* **7,** 439–450 (2009).
30. Snijder, E., Decroly, E. & Ziebuhr, J. in, 59–126 (2016).
31. Gosert, R., Kanjanahaluethai, A., Egger, D., Bienz, K. & Baker, S. C. RNA Replication of Mouse Hepatitis Virus Takes Place at Double-Membrane Vesicles. *Journal of Virology* **76,** 3697–3708 (2002).
32. Sawicki, S. G. & Sawicki, D. L. in, 499–506 (Adv Exp Med Biol ., 1995).
33. Zúñiga, S., Sola, I., Alonso, S. & Enjuanes, L. Sequence Motifs Involved in the Regulation of Discontinuous Coronavirus Subgenomic RNA Synthesis. *Journal of Virology* **78,** 980–994 (2004).
34. Malone, B., Urakova, N., Snijder, E. J. & Campbell, E. A. Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. *Nature Reviews Molecular Cell Biology* **23,** 21–39 (2022).
35. Hartenian, E. *et al.* The molecular virology of coronaviruses. *Journal of Biological Chemistry* **295,** 12910–12934 (2020).
36. Munster, V. J., Koopmans, M., van Doremalen, N., van Riel, D. & de Wit, E. A Novel Coronavirus Emerging in China — Key Questions for Impact Assessment. *New England Journal of Medicine* **382,** 692–694 (2020).
37. Wang, H. *et al.* Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *The Lancet* **399,** 1513–1536 (2022).
38. Liu, D. X., Liang, J. Q. & Fung, T. S. in *Encyclopedia of Virology* 428–440 (Elsevier, 2021).
39. CHAN-YEUNG, M. & XU, R.-H. SARS: epidemiology. *Respirology* **8,** S9–S14 (2003).
40. Van der Hoek, L. *et al.* Identification of a new human coronavirus. *Nature Medicine* **10,** 368–373 (2004).
41. Woo, P. C. Y. *et al.* Characterization and Complete Genome Sequence of a Novel Coronavirus, Coronavirus HKU1, from Patients with Pneumonia. *Journal of Virology* **79,** 884–895 (2005).
42. Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. & Fouchier, R. A. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *New England Journal of Medicine* **367,** 1814–1820 (2012).

43. Mulabbi, E. N., Tweyongyere, R. & Byarugaba, D. K. The history of the emergence and transmission of human coronaviruses. *Onderstepoort Journal of Veterinary Research* **88** (2021).
44. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* **382,** 727–733 (2020).
45. Wang, L., Byrum, B. & Zhang, Y. Detection and Genetic Characterization of Deltacoronavirus in Pigs, Ohio, USA, 2014. *Emerging Infectious Diseases* **20** (2014).
46. Pan, Y. *et al.* Discovery of a novel swine enteric alphacoronavirus (SeACoV) in southern China. *Veterinary Microbiology* **211,** 15–21 (2017).
47. Turlewicz-Podbielska, H. & Pomorska-Mól, M. Porcine Coronaviruses: Overview of the State of the Art. *Virologica Sinica* (2021).
48. Roe, C. K. & Alexander, T. J. A Disease of Nursing Pigs Previously Unreported in Ontario. *Canadian journal of comparative medicine and veterinary science* **22,** 305–7 (1958).
49. Greig, A. S. *et al.* A Hemagglutinating Virus Producing Encephalomyelitis in Baby Pigs. *Canadian journal of comparative medicine and veterinary science* **26,** 49–56 (1962).
50. Mora-Díaz, J. C., Piñeyro, P. E., Houston, E., Zimmerman, J. & Giménez-Lirola, L. G. Porcine Hemagglutinating Encephalomyelitis Virus: A Review. *Frontiers in Veterinary Science* **6** (2019).
51. Chen, Y. *et al.* Transmissible Gastroenteritis Virus: An Update Review and Perspective. *Viruses* **15,** 359 (2023).
52. Pensaert, M., Callebaut, P. & Vergote, J. Isolation of a porcine respiratory, non-enteric coronavirus related to transmissible gastroenteritis. *Veterinary Quarterly* **8** (1986).
53. Saif, L. J. & Jung, K. *Comparative pathogenesis of bovine and porcine respiratory coronaviruses in the animal host species and sars-cov-2 in humans* 2020.
54. Schultze, B. *et al.* Transmissible gastroenteritis coronavirus, but not the related porcine respiratory coronavirus, has a sialic acid (N-glycolylneuraminic acid) binding activity. *Journal of Virology* **70** (1996).
55. Yuan, D. *et al.* Isolation and Characterization of a Porcine Transmissible Gastroenteritis Coronavirus in Northeast China. *Frontiers in Veterinary Science* **8** (2021).
56. Chen, F., Knutson, T. P., Rossow, S., Saif, L. J. & Marthaler, D. G. Decline of transmissible gastroenteritis virus and its complex evolutionary relationship with porcine respiratory coronavirus in the United States. *Scientific Reports* (2019).
57. Wang, Q., Vlasova, A. N., Kenney, S. P. & Saif, L. J. Emerging and re-emerging coronaviruses in pigs. *Current Opinion in Virology* **34,** 39–49 (2019).

58. Zhang, H. *et al.* Prevalence, phylogenetic and evolutionary analysis of porcine deltacoronavirus in Henan province, China. *Preventive Veterinary Medicine* **166,** 8–15 (2019).
59. Zhou, L. *et al.* The re-emerging of SADS-CoV infection in pig herds in Southern China. *Transboundary and Emerging Diseases* **66,** 2180–2183 (2019).
60. Diamond, M. S. & Kanneganti, T.-D. Innate immunity: the first line of defense against SARS-CoV-2. *Nature Immunology* **23,** 165–176 (2022).
61. Kanneganti, T.-D. Intracellular innate immune receptors: Life inside the cell. *Immunological Reviews* **297,** 5–12 (2020).
62. Yoneyama, M. *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature Immunology* **5** (2004).
63. Seth, R. B., Sun, L., Ea, C.-K. & Chen, Z. J. Identification and Characterization of MAVS, a Mitochondrial Antiviral Signaling Protein that Activates NF-κB and IRF3. *Cell* **122** (2005).
64. Lin, R., Heylbroeck, C., Pitha, P. M. & Hiscott, J. Virus-Dependent Phosphorylation of the IRF-3 Transcription Factor Regulates Nuclear Translocation, Transactivation Potential, and Proteasome-Mediated Degradation. *Molecular and Cellular Biology* **18,** 2986–2996 (1998).
65. Katze, M. G., He, Y. & Gale, M. Viruses and interferon: a fight for supremacy. *Nature Reviews Immunology* **2,** 675–687 (2002).
66. Darnell, J., Kerr, I. & Stark, G. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264** (1994).
67. Schindler, C., Levy, D. E. & Decker, T. JAK-STAT Signaling: From Interferons to Cytokines. *Journal of Biological Chemistry* **282** (2007).
68. Schneider, W. M., Chevillotte, M. D. & Rice, C. M. Interferon-Stimulated Genes: A Complex Web of Host Defenses. *Annual Review of Immunology* **32,** 513–545 (2014).
69. Zhou, J.-h. *et al.* Type III Interferons in Viral Infection and Antiviral Immunity. *Cellular Physiology and Biochemistry* **51,** 173–185 (2018).
70. Bailey, C. C., Zhong, G., Huang, I.-C. & Farzan, M. IFITM-Family Proteins: The Cell's First Line of Antiviral Defense. *Annual Review of Virology* **1,** 261–283 (2014).
71. Brass, A. L. *et al.* The IFITM Proteins Mediate Cellular Resistance to Influenza A H1N1 Virus, West Nile Virus, and Dengue Virus. *Cell* **139,** 1243–1254 (2009).
72. Huang, I.-C. *et al.* Distinct Patterns of IFITM-Mediated Restriction of Filoviruses, SARS Coronavirus, and Influenza A Virus. *PLoS Pathogens* **7** (2011).
73. Chesarino, N. M. *et al.* <scp>IFITM</scp> 3 requires an amphipathic helix for antiviral activity. *EMBO reports* **18,** 1740–1751 (2017).

74. Liu, S.-Y. *et al.* Interferon-Inducible Cholesterol-25-Hydroxylase Broadly Inhibits Viral Entry by Production of 25-Hydroxycholesterol. *Immunity* **38,** 92–105 (2013).
75. Goujon, C. *et al.* Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* **502,** 559–562 (2013).
76. Schoggins, J. W. Interferon-Stimulated Genes: What Do They All Do? *Annual Review of Virology* **6,** 567–584 (2019).
77. Doyle, T. *et al.* The interferon-inducible isoform of NCOA7 inhibits endosome-mediated viral entry. *Nature Microbiology* **3,** 1369–1376 (2018).
78. Skorupka, K. A. *et al.* Hierarchical assembly governs TRIM5α recognition of HIV-1 and retroviral capsids. *Science Advances* **5** (2019).
79. Stremlau, M. *et al.* The cytoplasmic body component TRIM5α restricts HIV-1 infection in Old World monkeys. *Nature* **427,** 848–853 (2004).
80. Ganser-Pornillos, B. K. *et al.* Hexagonal assembly of a restricting TRIM5α protein. *Proceedings of the National Academy of Sciences* **108,** 534–539 (2011).
81. Haller, O., Staeheli, P., Schwemmle, M. & Kochs, G. Mx GTPases: dynamin-like antiviral machines of innate immunity. *Trends in Microbiology* **23,** 154–163 (2015).
82. Gariano, G. R. *et al.* The Intracellular DNA Sensor IFI16 Gene Acts as Restriction Factor for Human Cytomegalovirus Replication. *PLoS Pathogens* **8,** e1002498 (2012).
83. Batra, J. *et al.* Protein Interaction Mapping Identifies RBBP6 as a Negative Regulator of Ebola Virus Replication. *Cell* **175,** 1917–1930 (2018).
84. Chemudupati, M. *et al.* From APOBEC to ZAP: Diverse mechanisms used by cellular restriction factors to inhibit virus infections. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1866,** 382–394 (2019).
85. Zhang, H. *et al.* The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424,** 94–98 (2003).
86. Gizzi, A. S. *et al.* A naturally occurring antiviral ribonucleotide encoded by the human genome. *Nature* **558,** 610–614 (2018).
87. Schlee, M. & Hartmann, G. Discriminating self from non-self in nucleic acid sensing. *Nature Reviews Immunology* **16,** 566–580 (2016).
88. Abbas, Y. M., Pichlmair, A., Górna, M. W., Superti-Furga, G. & Nagar, B. Structural basis for viral 5-PPP-RNA recognition by human IFIT proteins. *Nature* **494,** 60–64 (2013).
89. Vladimer, G. I., GÃ³rna, M. W. & Superti-Furga, G. IFITs: Emerging Roles as Key Anti-Viral Proteins. *Frontiers in Immunology* **5** (2014).
90. Decroly, E., Ferron, F., Lescar, J. & Canard, B. Conventional and unconventional mechanisms for capping viral mRNA. *Nature Reviews Microbiology* **10,** 51–65 (2012).

91. Simmonds, P., Xia, W., Baillie, J. & McKinnon, K. Modelling mutational and selection pressures on dinucleotides in eukaryotic phyla –selection against CpG and UpA in cytoplasmically expressed RNA and in RNA viruses. *BMC Genomics* **14,** 610 (2013).
92. Di Giallonardo, F., Schlub, T. E., Shi, M. & Holmes, E. C. Dinucleotide Composition in Animal RNA Viruses Is Shaped More by Virus Family than by Host Species. *Journal of Virology* **91** (2017).
93. Guo, X., Ma, J., Sun, J. & Gao, G. The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. *Proceedings of the National Academy of Sciences* **104,** 151–156 (2007).
94. Guo, X., Carroll, J.-W. N., MacDonald, M. R., Goff, S. P. & Gao, G. The Zinc Finger Antiviral Protein Directly Binds to Specific Viral mRNAs through the CCCH Zinc Finger Motifs. *Journal of Virology* **78,** 12781–12787 (2004).
95. Zhou, A. Expression cloning of 2-5A-dependent RNAase: A uniquely regulated mediator of interferon action. *Cell* **72,** 753–765 (1993).
96. Zilberstein, A., Kimchi, A., Schmidt, A. & Revel, M. Isolation of two interferon-induced translational inhibitors: a protein kinase and an oligo-isoadenylate synthetase. *Proceedings of the National Academy of Sciences* **75,** 4734–4738 (1978).
97. George, C. X., John, L. & Samuel, C. E. An RNA Editor, Adenosine Deaminase Acting on Double-Stranded RNA (ADAR1). *Journal of Interferon & Cytokine Research* **34,** 437–446 (2014).
98. Perez-Caballero, D. *et al.* Tetherin Inhibits HIV-1 Release by Directly Tethering Virions to Cells. *Cell* **139,** 499–511 (2009).
99. Samuel, C. E. Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. *Virology* **411,** 180–193 (2011).
100. Li, M. M. H. *et al.* Characterization of Novel Splice Variants of Zinc Finger Antiviral Protein (ZAP). *Journal of Virology* **93** (2019).
101. Liu, J., Qian, C. & Cao, X. Post-Translational Modification Control of Innate Immunity. *Immunity* **45,** 15–30 (2016).
102. Schulz, O. *et al.* Protein Kinase R Contributes to Immunity against Specific Viruses by Regulating Interferon mRNA Integrity. *Cell Host & Microbe* **7,** 354–361 (2010).
103. Anderson, P. Post-transcriptional control of cytokine production. *Nature Immunology* **9,** 353–359 (2008).
104. Van der Made, C. I. *et al.* Presence of Genetic Variants Among Young Men With Severe COVID-19. *JAMA* **324,** 663 (2020).
105. Wickenhagen, A. *et al.* A prenylated dsRNA sensor protects against severe COVID-19. *Science* **374** (2021).
106. Zhang, Q. *et al.* Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **370** (2020).

107. Martin-Sancho, L. *et al.* Functional landscape of SARS-CoV-2 cellular restriction. *Molecular Cell* **81** (2021).
108. Lokugamage, K. G. *et al.* Type I Interferon Susceptibility Distinguishes SARS-CoV-2 from SARS-CoV. *Journal of Virology* **94** (2020).
109. Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* **472,** 481–485 (2011).
110. Mar, K. B. *et al.* LY6E mediates an evolutionarily conserved enhancement of virus infection by targeting a late entry step. *Nature Communications* **9** (2018).
111. Pfaender, S. *et al.* LY6E impairs coronavirus fusion and confers immune control of viral disease. *Nature Microbiology* **5** (2020).
112. Mar, K. B. *et al.* LY6E is a pan-coronavirus restriction factor in the respiratory tract. *Nature Microbiology* **8,** 1587–1599 (2023).
113. Karki, R. & Kanneganti, T.-D. The 'cytokine storm': molecular mechanisms and therapeutic prospects. *Trends in Immunology* **42,** 681–705 (2021).
114. Ramasamy, S. & Subbian, S. Critical Determinants of Cytokine Storm and Type I Interferon Response in COVID-19 Pathogenesis. *Clinical Microbiology Reviews* **34** (2021).
115. Ribero, M. S., Jouvenet, N., Dreux, M. & Nisole, S. *Interplay between SARS-CoV-2 and the type I interferon response* 2020.
116. Channappanavar, R. *et al.* Dysregulated Type I Interferon and Inflammatory Monocyte-Macrophage Responses Cause Lethal Pneumonia in SARS-CoV-Infected Mice. *Cell Host & Microbe* **19** (2016).
117. Van Valen, L. A new evolutionary law. *Evolutionary Theory* **1,** 1–30 (1973).
118. García-Sastre, A. Ten Strategies of Interferon Evasion by Viruses. *Cell host & microbe* **22,** 176–184 (2017).
119. Beachboard, D. C. & Horner, S. M. Innate immune evasion strategies of DNA and RNA viruses. *Current Opinion in Microbiology* **32,** 113–119 (2016).
120. Neufeldt, C. J. *et al.* The Hepatitis C Virus-Induced Membranous Web and Associated Nuclear Transport Machinery Limit Access of Pattern Recognition Receptors to Viral Replication Sites. *PLOS Pathogens* **12,** e1005428 (2016).
121. Angelini, M. M., Akhlaghpour, M., Neuman, B. W. & Buchmeier, M. J. Severe Acute Respiratory Syndrome Coronavirus Nonstructural Proteins 3, 4, and 6 Induce Double-Membrane Vesicles. *mBio* **4** (2013).
122. Oudshoorn, D. *et al.* Expression and Cleavage of Middle East Respiratory Syndrome Coronavirus nsp3-4 Polyprotein Induce the Formation of Double-Membrane Vesicles That Mimic Those Associated with Coronaviral RNA Replication. *mBio* **8** (2017).
123. Wolff, G. *et al.* A molecular pore spans the double membrane of the coronavirus replication organelle. *Science* **369,** 1395–1398 (2020).

124. Lahaye, X. *et al.* The Capsids of HIV-1 and HIV-2 Determine Immune Detection of the Viral cDNA by the Innate Sensor cGAS in Dendritic Cells. *Immunity* **39,** 1132–1142 (2013).
125. Feng, Q. *et al.* Enterovirus 2A ^{pro} Targets MDA5 and MAVS in Infected Cells. *Journal of Virology* **88,** 3369–3378 (2014).
126. Liu, Y. *et al.* SARS-CoV-2 Nsp5 Demonstrates Two Distinct Mechanisms Targeting RIG-I and MAVS To Evade the Innate Immune Response. *mBio* **12** (2021).
127. Liu, G. *et al.* ISG15-dependent activation of the sensor MDA5 is antagonized by the SARS-CoV-2 papain-like protease to evade host innate immunity. *Nature Microbiology* **6,** 467–478 (2021).
128. Zhao, X. *et al.* Interferon Control of Human Coronavirus Infection and Viral Evasion: Mechanistic Insights and Implications for Antiviral Drug and Vaccine Development. *Journal of Molecular Biology* **434,** 167438 (2022).
129. Xia, H. *et al.* Evasion of Type I Interferon by SARS-CoV-2. *Cell Reports* **33,** 108234 (2020).
130. Lau, L., Gray, E. E., Brunette, R. L. & Stetson, D. B. DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway. *Science* **350,** 568–571 (2015).
131. Ma, Z. *et al.* Modulation of the cGAS-STING DNA sensing pathway by gammaherpesviruses. *Proceedings of the National Academy of Sciences* **112** (2015).
132. Maringer, K. & Fernandez-Sesma, A. Message in a bottle: lessons learned from antagonism of STING signalling during RNA virus infection. *Cytokine & Growth Factor Reviews* **25,** 669–679 (2014).
133. West, A. P. *et al.* Mitochondrial DNA stress primes the antiviral innate immune response. *Nature* **520,** 553–557 (2015).
134. Huang, C. *et al.* SARS Coronavirus nsp1 Protein Induces Template-Dependent Endonucleolytic Cleavage of mRNAs: Viral mRNAs Are Resistant to nsp1-Induced RNA Cleavage. *PLoS Pathogens* **7,** e1002433 (2011).
135. Khaperskyy, D. A., Schmaling, S., Larkins-Ford, J., McCormick, C. & Gaglia, M. M. Selective Degradation of Host RNA Polymerase II Transcripts by Influenza A Virus PA-X Host Shutoff Protein. *PLOS Pathogens* **12,** e1005427 (2016).
136. Gaucherand, L. *et al.* The Influenza A Virus Endoribonuclease PA-X Usurps Host mRNA Processing Machinery to Limit Host Gene Expression. *Cell Reports* **27,** 776–792 (2019).
137. Shu, M., Taddeo, B., Zhang, W. & Roizman, B. Selective degradation of mRNAs by the HSV host shutoff RNase is regulated by the U _L 47 tegument protein. *Proceedings of the National Academy of Sciences* **110** (2013).
138. Lam, Q. *et al.* Herpes simplex virus VP16 rescues viral mRNA from destruction by the virion host shutoff function. *The EMBO journal* **15,** 2575–81 (1996).

139. Taddeo, B., Sciortino, M. T., Zhang, W. & Roizman, B. Interaction of herpes simplex virus RNase with VP16 and VP22 is required for the accumulation of the protein but not for accumulation of mRNA. *Proceedings of the National Academy of Sciences* **104,** 12163–12168 (2007).
140. Guha, D. & Ayyavoo, V. Innate Immune Evasion Strategies by Human Immunodeficiency Virus Type 1. *ISRN AIDS* **2013,** 1–10 (2013).
141. Neil, S. J. D., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451,** 425–430 (2008).
142. McNatt, M. W., Zang, T. & Bieniasz, P. D. Vpu Binds Directly to Tetherin and Displaces It from Nascent Virions. *PLoS Pathogens* **9,** e1003299 (2013).
143. Hofmann, H. *et al.* The Vpx Lentiviral Accessory Protein Targets SAMHD1 for Degradation in the Nucleus. *Journal of Virology* **86,** 12552–12560 (2012).
144. Zeng, C. *et al.* SERINC proteins potentiate antiviral type I IFN production and proinflammatory signaling pathways. *Science Signaling* **14** (2021).
145. Rosa, A. *et al.* HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. *Nature* **526,** 212–217 (2015).
146. Odon, V., Fiddaman, S. R., Smith, A. L. & Simmonds, P. Comparison of CpG- and UpA-mediated restriction of RNA virus replication in mammalian and avian cells and investigation of potential ZAP-mediated shaping of host transcriptome compositions. *RNA* **28,** 1089–1109 (2022).
147. Shi, G. *et al.* Opposing activities of IFITM proteins in SARS-CoV-2 infection. *The EMBO Journal* **40** (2021).
148. Prelli Bozzo, C. *et al.* IFITM proteins promote SARS-CoV-2 infection and are targets for virus inhibition in vitro. *Nature Communications* **12,** 4584 (2021).
149. Zhao, X. *et al.* Interferon induction of IFITM proteins promotes infection by human coronavirus OC43. *Proceedings of the National Academy of Sciences* **111,** 6756–6761 (2014).
150. Seo, J.-Y., Yaneva, R., Hinson, E. R. & Cresswell, P. Human Cytomegalovirus Directly Induces the Antiviral Protein Viperin to Enhance Infectivity. *Science* **332,** 1093–1097 (2011).
151. Viswanathan, K. *et al.* BST2/Tetherin Enhances Entry of Human Cytomegalovirus. *PLoS Pathogens* **7,** e1002332 (2011).
152. Schindewolf, C. *et al.* SARS-CoV-2 Uses Nonstructural Protein 16 To Evade Restriction by IFIT1 and IFIT3. *Journal of Virology* **97** (2023).
153. Karesh, W. B. *et al.* Ecology of zoonoses: natural and unnatural histories. *The Lancet* **380,** 1936–1945 (2012).
154. Shaw, A. E. *et al.* Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. *PLoS Biology* (2017).

155. Long, J. S., Mistry, B., Haslam, S. M. & Barclay, W. S. Host and viral determinants of influenza A virus species specificity. *Nature Reviews Microbiology* **17,** 67–81 (2019).
156. Hjelle, B. & Glass, G. E. Outbreak of Hantavirus Infection in the Four Corners Region of the United States in the Wake of the 1997–1998 El Niño–Southern Oscillation. *The Journal of Infectious Diseases* **181,** 1569–1573 (2000).
157. Ducatez, M., Webster, R. & Webby, R. Animal influenza epidemiology. *Vaccine* **26,** D67–D69 (2008).
158. Rupprecht, C. E., Hanlon, C. A. & Hemachudha, T. Rabies re-examined. *The Lancet Infectious Diseases* **2,** 327–343 (2002).
159. Plowright, R. K. *et al.* Pathways to zoonotic spillover. *Nature Reviews Microbiology* **15,** 502–510 (2017).
160. Weber, T. P. & Stilianakis, N. I. Inactivation of influenza A viruses in the environment and modes of transmission: A critical review. *Journal of Infection* **57,** 361–373 (2008).
161. Boone, S. A. & Gerba, C. P. Significance of Fomites in the Spread of Respiratory and Enteric Viral Disease. *Applied and Environmental Microbiology* **73,** 1687–1696 (2007).
162. Peros, C. S., Dasgupta, R., Kumar, P. & Johnson, B. A. Bushmeat, wet markets, and the risks of pandemics: Exploring the nexus through systematic review of scientific disclosures. *Environmental Science & Policy* **124,** 1–11 (2021).
163. Lam, S. K. & Chua, K. B. Nipah Virus Encephalitis Outbreak in Malaysia. *Clinical Infectious Diseases* **34,** S48–S51 (2002).
164. Woo, P. C., Lau, S. K. & Yuen, K.-y. Infectious diseases emerging from Chinese wet-markets: zoonotic origins of severe respiratory viral infections. *Current Opinion in Infectious Diseases* **19,** 401–407 (2006).
165. Perdue, M. L. & Swayne, D. E. Public Health Risk from Avian Influenza Viruses. *Avian Diseases* **49,** 317–327 (2005).
166. Tatem, A. J., Hay, S. I. & Rogers, D. J. Global traffic and disease vector dispersal. *Proceedings of the National Academy of Sciences* **103,** 6242–6247 (2006).
167. Bartlow, A. W. *et al.* Forecasting Zoonotic Infectious Disease Response to Climate Change: Mosquito Vectors and a Changing Environment. *Veterinary Sciences* **6,** 40 (2019).
168. Lowen, A. C. & Steel, J. Roles of Humidity and Temperature in Shaping Influenza Seasonality. *Journal of Virology* **88,** 7692–7695 (2014).
169. Chan, K. H. *et al.* The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus. *Advances in Virology* **2011,** 1–7 (2011).

170. Casanova, L. M., Jeon, S., Rutala, W. A., Weber, D. J. & Sobsey, M. D. Effects of Air Temperature and Relative Humidity on Coronavirus Survival on Surfaces. *Applied and Environmental Microbiology* **76,** 2712–2717 (2010).
171. Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS Pandemic. *Cold Spring Harbor Perspectives in Medicine* **1,** a006841–a006841 (2011).
172. Sauter, D. & Kirchhoff, F. Key Viral Adaptations Preceding the AIDS Pandemic. *Cell Host & Microbe* **25,** 27–38 (2019).
173. Sauter, D. *et al.* Tetherin-Driven Adaptation of Vpu and Nef Function and the Evolution of Pandemic and Nonpandemic HIV-1 Strains. *Cell Host & Microbe* **6,** 409–421 (2009).
174. Zuliani-Alvarez, L. *et al.* Evasion of cGAS and TRIM5 defines pandemic HIV. *Nature Microbiology* **7,** 1762–1776 (2022).
175. Rogers, G. N. & Paulson, J. C. Receptor determinants of human and animal influenza virus isolates: Differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology* **127,** 361–373 (1983).
176. Gagneux, P. *et al.* Human-specific Regulation of α2–6-linked Sialic Acids. *Journal of Biological Chemistry* **278,** 48245–48250 (2003).
177. Russell, C. J., Hu, M. & Okda, F. A. Influenza Hemagglutinin Protein Stability, Activation, and Pandemic Risk. *Trends in Microbiology* **26,** 841–853 (2018).
178. Massin, P., van der Werf, S. & Naffakh, N. Residue 627 of PB2 Is a Determinant of Cold Sensitivity in RNA Replication of Avian Influenza Viruses. *Journal of Virology* **75,** 5398–5404 (2001).
179. Scull, M. A. *et al.* Avian Influenza Virus Glycoproteins Restrict Virus Replication and Spread through Human Airway Epithelium at Temperatures of the Proximal Airways. *PLoS Pathogens* **5,** e1000424 (2009).
180. Long, J. S. *et al.* Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature* **529,** 101–104 (2016).
181. Pinto, R. M. *et al.* BTN3A3 evasion promotes the zoonotic potential of influenza A viruses. *Nature* **619,** 338–347 (2023).
182. Sang, Y., Bergkamp, J. & Blecha, F. Molecular Evolution of the Porcine Type I Interferon Family: Subtype-Specific Expression and Antiviral Activity. *PLoS ONE* **9,** e112378 (2014).
183. Levy, D. E., Marié, I. J. & Durbin, J. E. Induction and function of type I and III interferon in response to viral infection. *Current Opinion in Virology* **1,** 476–486 (2011).
184. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nature Reviews Immunology* **14,** 36–49 (2014).
185. Sanjuán, R. & Domingo-Calap, P. Mechanisms of viral mutation. *Cellular and Molecular Life Sciences* **73** (2016).

186. Simon-Loriere, E. & Holmes, E. C. Why do RNA viruses recombine? *Nature Reviews Microbiology* **9,** 617–626 (2011).
187. Ma, W., Kahn, R. E. & Richt, J. A. The pig as a mixing vessel for influenza viruses: Human and veterinary implications. *Journal of molecular and genetic medicine : an international journal of biomedical research* **3,** 158–66 (2008).
188. Lai, M. M. RNA recombination in animal and plant viruses. *Microbiological Reviews* **56,** 61–79 (1992).
189. Chare, E. R., Gould, E. A. & Holmes, E. C. Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *Journal of General Virology* **84,** 2691–2703 (2003).
190. Kawasaki, Y., Abe, H. & Yasuda, J. Comparison of genome replication fidelity between SARS-CoV-2 and influenza A virus in cell culture. *Scientific Reports* **13,** 13105 (2023).
191. Wells, H. L. *et al.* The coronavirus recombination pathway. *Cell Host & Microbe* **31,** 874–889 (2023).
192. Graham, R. L. & Baric, R. S. Recombination, Reservoirs, and the Modular Spike: Mechanisms of Coronavirus Cross-Species Transmission. *Journal of Virology* **84** (2010).
193. Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiology* **5,** 562–569 (2020).
194. Wells, H. L. *et al.* The evolutionary history of ACE2 usage within the coronavirus subgenus Sarbecovirus. *Virus Evolution* **7** (2021).
195. Anthony, S. J. *et al.* Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus. *mBio* **8** (2017).
196. Lau, S. K. P. *et al.* Molecular Epidemiology of Human Coronavirus OC43 Reveals Evolution of Different Genotypes over Time and Recent Emergence of a Novel Genotype due to Natural Recombination. *Journal of Virology* **85,** 11325–11337 (2011).
197. Tao, Y. *et al.* Surveillance of Bat Coronaviruses in Kenya Identifies Relatives of Human Coronaviruses NL63 and 229E and Their Recombination History. *Journal of Virology* **91** (2017).
198. Woo, P. C. Y. *et al.* Phylogenetic and recombination analysis of coronavirus HKU1, a novel coronavirus from patients with pneumonia. *Archives of Virology* **150,** 2299–2311 (2005).
199. Chen, D.-Y. *et al.* Spike and nsp6 are key determinants of SARS-CoV-2 Omicron BA.1 attenuation. *Nature* **615,** 143–150 (2023).
200. Lorusso, A. *et al.* Gain, Preservation, and Loss of a Group 1a Coronavirus Accessory Glycoprotein. *Journal of Virology* **82,** 10312–10317 (2008).

201. Le Poder, S. Feline and Canine Coronaviruses: Common Genetic and Pathobiological Features. *Advances in Virology* **2011,** 1–11 (2011).
202. Jaimes, J. A., Millet, J. K., Stout, A. E., André, N. M. & Whittaker, G. R. A Tale of Two Viruses: The Distinct Spike Glycoproteins of Feline Coronaviruses. *Viruses* **12,** 83 (2020).
203. Attipa, C. *et al.* Emergence and spread of feline infectious peritonitis due to a highly pathogenic canine/feline recombinant coronavirus. *bioRxiv,* 2023.11.08.566182 (2023).
204. Boniotti, M. B. *et al.* Porcine Epidemic Diarrhea Virus and Discovery of a Recombinant Swine Enteric Coronavirus, Italy. *Emerging Infectious Diseases* **22,** 83–87 (2016).
205. Ruiz-Aravena, M. *et al.* Ecology, evolution and spillover of coronaviruses from bats. *Nature Reviews Microbiology* **20,** 299–314 (2022).
206. Wang, Q. *et al.* Tracing the origins of SARS-CoV-2: lessons learned from the past. *Cell Research* **31,** 1139–1141 (2021).
207. Dhama, K. *et al.* SARS-CoV-2 jumping the species barrier: Zoonotic lessons from SARS, MERS and recent advances to combat this pandemic virus. *Travel Medicine and Infectious Disease* **37,** 101830 (2020).
208. Mair, K. *et al.* The porcine innate immune system: An update. *Developmental & Comparative Immunology* **45,** 321–343 (2014).
209. Dawson, H. D., Smith, A. D., Chen, C. & Urban, J. F. An in-depth comparison of the porcine, murine and human inflammasomes; lessons from the porcine genome and transcriptome. *Veterinary Microbiology* **202,** 2–15 (2017).
210. Suschak, J. J., Wang, S., Fitzgerald, K. A. & Lu, S. Identification of Aim2 as a Sensor for DNA Vaccines. *The Journal of Immunology* **194,** 630–636 (2015).
211. Zhu, J., Lai, K., Brownile, R., Babiuk, L. A. & Mutwiri, G. K. Porcine TLR8 and TLR7 are both activated by a selective TLR7 ligand, imiquimod. *Molecular Immunology* **45,** 3238–3243 (2008).
212. Zhang, X. *et al.* Guanylate-binding protein 1 inhibits nuclear delivery of pseudorabies virus by disrupting structure of actin filaments. *Veterinary Research* **54,** 21 (2023).
213. Muñoz-Moreno, R. *et al.* Antiviral Role of IFITM Proteins in African Swine Fever Virus Infection. *PLOS ONE* **11,** e0154366 (2016).
214. Zhou, J. *et al.* Porcine Mx1 Protein Inhibits Classical Swine Fever Virus Replication by Targeting Nonstructural Protein NS5B. *Journal of Virology* **92** (2018).
215. Li, C. *et al.* Antiviral activity of ISG15 against classical swine fever virus replication in porcine alveolar macrophages via inhibition of autophagy by ISGylating BECN1. *Veterinary Research* **51,** 22 (2020).

216. Zhang, Y., Burke, C. W., Ryman, K. D. & Klimstra, W. B. Identification and Characterization of Interferon-Induced Proteins That Inhibit Alphavirus Replication. *Journal of Virology* **81,** 11246–11255 (2007).
217. Itsui, Y. *et al.* Expressional screening of interferon-stimulated genes for antiviral activity against hepatitis C virus replication. *Journal of Viral Hepatitis* **13,** 690–700 (2006).
218. Meurens, F., Summerfield, A., Nauwynck, H., Saif, L. & Gerdts, V. The pig: a model for human infectious diseases. *Trends in Microbiology* **20,** 50–57 (2012).
219. Edwards, C. E. *et al.* Swine acute diarrhea syndrome coronavirus replication in primary human cells reveals potential susceptibility to infection. *Proceedings of the National Academy of Sciences* (2020).
220. Lednicky, J. A. *et al.* Independent infections of porcine deltacoronavirus among Haitian children. *Nature* **600,** 133–137 (2021).
221. Dittmann, M. *et al.* A Serpin Shapes the Extracellular Environment to Prevent Influenza A Virus Maturation. *Cell* **160,** 631–643 (2015).
222. Opriessnig, T. & Huang, Y.-W. SARS-CoV-2 does not infect pigs, but this has to be verified regularly. *Xenotransplantation* **29** (2022).
223. Zhang, X. *et al.* Cytokine Responses in Porcine Respiratory Coronavirus-Infected Pigs Treated with Corticosteroids as a Model for Severe Acute Respiratory Syndrome. *Journal of Virology* **82** (2008).
224. Jung, K. *et al.* Altered Pathogenesis of Porcine Respiratory Coronavirus in Pigs due to Immunosuppressive Effects of Dexamethasone: Implications for Corticosteroid Use in Treatment of Severe Acute Respiratory Syndrome Coronavirus. *Journal of Virology* (2007).
225. Daniels, A. *et al.* One for all—human kidney Caki-1 cells are highly susceptible to infection with corona- and other respiratory viruses. *Journal of Virology* **97** (2023).
226. Craig, N. *et al.* Direct Lysis RT-qPCR of SARS-CoV-2 in Cell Culture Supernatant Allows for Fast and Accurate Quantification. *Viruses* **14,** 508 (2022).
227. McCray, P. B. *et al.* Lethal Infection of K18- hACE2 Mice Infected with Severe Acute Respiratory Syndrome Coronavirus. *Journal of Virology* **81,** 813–821 (2007).
228. Whitelaw, C. A. *et al.* Efficient generation of transgenic pigs using equine infectious anaemia virus (EIAV) derived vector. *FEBS Letters* **571,** 233–236 (2004).
229. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Scientific Reports* **7,** 16878 (2017).
230. Ram, S. *et al.* Pixelwise H-score: A novel digital image analysis-based metric to quantify membrane biomarker expression from immunohistochemistry images. *PLOS ONE* **16,** e0245638 (2021).

231. Ogando, N. S. *et al.* SARS-coronavirus-2 replication in Vero E6 cells: replication kinetics, rapid adaptation and cytopathology. *Journal of General Virology* **101,** 925–940 (2020).
232. Yim-im, W. *et al.* Comparison of ZMAC and MARC-145 Cell Lines for Improving Porcine Reproductive and Respiratory Syndrome Virus Isolation from Clinical Samples. *Journal of Clinical Microbiology* **59** (2021).
233. Zhang, Y. *et al.* Porcine Epidemic Diarrhea Virus: An Updated Overview of Virus Epidemiology, Virulence Variation Patterns and Virus–Host Interactions. *Viruses* **14,** 2434 (2022).
234. Wu, Y. *et al.* Characterization and pathogenicity of Vero cell-attenuated porcine epidemic diarrhea virus CT strain. *Virology Journal* **16,** 121 (2019).
235. Glover, D. J., Lipps, H. J. & Jans, D. A. Towards safe, non-viral therapeutic gene expression in humans. *Nature Reviews Genetics* **6,** 299–310 (2005).
236. Vorburger, S. A. & Hunt, K. K. Adenoviral Gene Therapy. *The Oncologist* **7,** 46–59 (2002).
237. Mann, R., Mulligan, R. C. & Baltimore, D. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33,** 153–159 (1983).
238. Dufait, I. *et al.* Retroviral and Lentiviral Vectors for the Induction of Immunological Tolerance. *Scientifica* **2012,** 1–14 (2012).
239. Naldini, L. *et al.* In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. *Science* **272,** 263–267 (1996).
240. Addgene. Lentiviral Guide. *https://www.addgene.org/guides/lentivirus/.*
241. Durand, S. & Cimarelli, A. The Inside Out of Lentiviral Vectors. *Viruses* **3,** 132–159 (2011).
242. Ikeda, Y., Collins, M., Radcliffe, P., Mitrophanous, K. & Takeuchi, Y. Gene transduction efficiency in cells of different species by HIV and EIAV vectors. *Gene Therapy* **9,** 932–938 (2002).
243. Chen, X., He, J. & Chang, L.-J. Alteration of T cell immunity by lentiviral transduction of human monocyte-derived dendritic cells. *Retrovirology* **1,** 37 (2004).
244. Tan, P. H. *et al.* Modulation of human dendritic-cell function following transduction with viral vectors: implications for gene therapy. *Blood* **105,** 3824–3832 (2005).
245. Mitrophanous, K. A. *et al.* Stable gene transfer to the nervous system using a non-primate lentiviral vector. *Gene Therapy* **6,** 1808–1818 (1999).
246. Johnston, J. C. *et al.* Minimum Requirements for Efficient Transduction of Dividing and Nondividing Cells by Feline Immunodeficiency Virus Vectors. *Journal of Virology* **73,** 4991–5000 (1999).

247. Matukonis, M. *et al.* Development of Second- and Third-Generation Bovine Immunodeficiency Virus-Based Gene Transfer Systems. *Human Gene Therapy* **13,** 1293–1303 (2002).
248. Tresnan, D. B., Levis, R. & Holmes, K. V. Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *Journal of Virology* **70,** 8669–8674 (1996).
249. Kane, M. *et al.* Identification of Interferon-Stimulated Genes with Antiretroviral Activity. *Cell Host & Microbe* **20,** 392–405 (2016).
250. Lin, Y.-T. *et al.* Human cytomegalovirus evades ZAP detection by suppressing CpG dinucleotides in the major immediate early 1 gene. *PLOS Pathogens* **16,** e1008844 (2020).
251. Wang, B. *et al.* Porcine Deltacoronavirus Engages the Transmissible Gastroenteritis Virus Functional Receptor Porcine Aminopeptidase N for Infectious Cellular Entry. *Journal of Virology* **92** (2018).
252. Li, W. *et al.* Broad receptor engagement of an emerging global coronavirus may potentiate its diverse cross-species transmissibility. *Proceedings of the National Academy of Sciences* **115** (2018).
253. Wu, L. *et al.* Broad host range of SARS-CoV-2 and the molecular basis for SARS-CoV-2 binding to cat ACE2. *Cell Discovery* **6,** 68 (2020).
254. Keep, S. *et al.* Porcine Respiratory Coronavirus as a Model for Acute Respiratory Coronavirus Disease. *Frontiers in Immunology* **13** (2022).
255. Cameroni, E. *et al.* Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* **602,** 664–670 (2022).
256. Cui, Z. *et al.* Structural and functional characterizations of infectivity and immune evasion of SARS-CoV-2 Omicron. *Cell* **185,** 860–871 (2022).
257. Lan, J. *et al.* Structural insights into the SARS-CoV-2 Omicron RBD-ACE2 interaction. *Cell Research* **32,** 593–595 (2022).
258. Li, L. *et al.* Structural basis of human ACE2 higher binding affinity to currently circulating Omicron SARS-CoV-2 sub-variants BA.2 and BA.1.1. *Cell* **185,** 2952–2960 (2022).
259. Thakur, N. *et al.* SARS-CoV-2 variants of concern alpha, beta, gamma and delta have extended ACE2 receptor host ranges. *Journal of General Virology* **103** (2022).
260. Bouhaddou, M. *et al.* SARS-CoV-2 variants evolve convergent strategies to remodel the host response. *Cell* **186,** 4597–4614 (2023).
261. Thorne, L. G. *et al.* Evolution of enhanced innate immune evasion by SARS-CoV-2. *Nature* **602,** 487–495 (2022).
262. Wei, C. *et al.* Evidence for a mouse origin of the SARS-CoV-2 Omicron variant. *Journal of Genetics and Genomics* **48,** 1111–1121 (2021).

263. Mallapaty, S. Where did Omicron come from? Three key theories. *Nature* **602,** 26–28 (2022).
264. Willett, B. J. *et al.* SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway. *Nature Microbiology* **7,** 1161–1179 (2022).
265. Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nature Reviews Molecular Cell Biology* **23,** 3–20 (2022).
266. Meng, B. *et al.* Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity. *Nature* **603,** 706–714 (2022).
267. Van Boxel-Dezaire, A. H., Rani, M. S. & Stark, G. R. Complex Modulation of Cell Type-Specific Signaling in Response to Type I Interferons. *Immunity* **25,** 361–372 (2006).
268. Muckenhuber, M. *et al.* Epigenetic signals that direct cell type–specific interferon beta response in mouse cells. *Life Science Alliance* **6,** e202201823 (2023).
269. Zhou, H. *et al.* Genome-Scale RNAi Screen for Host Factors Required for HIV Replication. *Cell Host & Microbe* **4,** 495–504 (2008).
270. Brass, A. L. *et al.* Identification of Host Proteins Required for HIV Infection Through a Functional Genomic Screen. *Science* **319,** 921–926 (2008).
271. Karlas, A. *et al.* Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* **463,** 818–822 (2010).
272. Li, Q. *et al.* A genome-wide genetic screen for host factors required for hepatitis C virus propagation. *Proceedings of the National Academy of Sciences* **106,** 16410–16415 (2009).
273. Zhu, Y. *et al.* A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. *Nature Communications* **12,** 961 (2021).
274. Li, Y. *et al.* Genome-wide CRISPR screen for Zika virus resistance in human neural cells. *Proceedings of the National Academy of Sciences* **116,** 9527–9532 (2019).
275. Song, Y. *et al.* A genome-wide CRISPR/Cas9 gene knockout screen identifies immunoglobulin superfamily DCC subclass member 4 as a key host factor that promotes influenza virus endocytosis. *PLOS Pathogens* **17,** e1010141 (2021).
276. Wei, J. *et al.* Genome-wide CRISPR Screens Reveal Host Factors Critical for SARS-CoV-2 Infection. *Cell* **184,** 76–91 (2021).
277. Jones, C. E., Tan, W. S., Grey, F. & Hughes, D. J. Discovering antiviral restriction factors and pathways using genetic screens. *Journal of General Virology* **102** (2021).
278. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159,** 647–661 (2014).
279. Chavez, A. *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nature Methods* **12,** 326–328 (2015).

280. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517,** 583–588 (2015).
281. Danziger, O., Patel, R. S., DeGrace, E. J., Rosen, M. R. & Rosenberg, B. R. Inducible CRISPR activation screen for interferon-stimulated genes identifies OAS1 as a SARS-CoV-2 restriction factor. *PLOS Pathogens* **18,** e1010464 (2022).
282. Arnoldo, A. *et al.* A genome scale overexpression screen to reveal drug activity in human cells. *Genome Medicine* **6,** 32 (2014).
283. Lee, W. J. *et al.* Genome-Wide Overexpression Screen Identifies Genes Able to Bypass p16-Mediated Senescence in Melanoma. *SLAS Discovery* **22,** 298–308 (2017).
284. Wu, J. *et al.* Cyclic GMP-AMP Is an Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA. *Science* **339,** 826–830 (2013).
285. Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway. *Science* **339,** 786–791 (2013).
286. Abrams, M. E. *et al.* Oxysterols provide innate immunity to bacterial infection by mobilizing cell surface accessible cholesterol. *Nature Microbiology* **5,** 929–942 (2020).
287. Rinkenberger, N. *et al.* Overexpression screen of interferon-stimulated genes identifies RARRES3 as a restrictor of Toxoplasma gondii infection. *eLife* **10** (2021).
288. Schoggins, J. W. *et al.* Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity. *Nature* **505,** 691–695 (2014).
289. Kuroda, M. *et al.* Identification of interferon-stimulated genes that attenuate Ebola virus infection. *Nature Communications* **11,** 2953 (2020).
290. Feng, J. *et al.* Interferon-Stimulated Gene (ISG)-Expression Screening Reveals the Specific Antibunyaviral Activity of ISG20. *Journal of Virology* **92** (2018).
291. McDougal, M. B. *et al.* Interferon inhibits a model <scp>RNA</scp> virus via a limited set of inducible effector genes. *EMBO reports* **24** (2023).
292. Seifert, L. L. *et al.* The ETS transcription factor ELF1 regulates a broadly antiviral program distinct from the type I interferon response. *PLOS Pathogens* **15,** e1007634 (2019).
293. Sironi, M., Cagliani, R., Forni, D. & Clerici, M. Evolutionary insights into host–pathogen interactions from mammalian sequence data. *Nature Reviews Genetics* **16,** 224–236 (2015).
294. Mahgoub, M. *et al.* Dual histone methyl reader ZCWPW1 facilitates repair of meiotic double strand breaks in male mice. *eLife* **9** (2020).
295. Wells, D. *et al.* ZCWPW1 is recruited to recombination hotspots by PRDM9 and is essential for meiotic double strand break repair. *eLife* **9** (2020).

296. Huang, T. *et al.* The histone modification reader ZCWPW1 links histone methylation to PRDM9-induced double-strand break repair. *eLife* **9** (2020).
297. Dang, W. *et al.* Establishment of a CRISPR/Cas9 knockout library for screening type I interferon-inducible antiviral effectors in pig cells. *Frontiers in Immunology* **13** (2022).
298. Zhou, P. *et al.* Molecular Characterization of Transcriptome-wide Interactions between Highly Pathogenic Porcine Reproductive and Respiratory Syndrome Virus and Porcine Alveolar Macrophages in vivo. *International Journal of Biological Sciences* **7,** 947–959 (2011).
299. Gaidt, M. M. *et al.* Self-guarding of MORC3 enables virulence factor-triggered immunity. *Nature* **600,** 138–142 (2021).
300. Ptaschinski, C. *et al.* RSV-Induced H3K4 Demethylase KDM5B Leads to Regulation of Dendritic Cell-Derived Innate Cytokines and Exacerbates Pathogenesis In Vivo. *PLoS pathogens* **11,** e1004978 (2015).
301. Rusinova, I. *et al.* INTERFEROME v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Research* **41,** D1040–D1046 (2012).
302. Rebendenne, A. *et al.* Bidirectional genome-wide CRISPR screens reveal host factors regulating SARS-CoV-2, MERS-CoV and seasonal HCoVs. *Nature Genetics* **54,** 1090–1102 (2022).
303. Chakrabarti, A., Jha, B. K. & Silverman, R. H. New Insights into the Role of RNase L in Innate Immunity. *Journal of Interferon & Cytokine Research* **31,** 49–57 (2011).
304. Ficarelli, M. *et al.* KHNYN is essential for the zinc finger antiviral protein (ZAP) to restrict HIV-1 containing clustered CpG dinucleotides. *eLife* **8** (2019).
305. Zheng, X. *et al.* TRIM25 Is Required for the Antiviral Activity of Zinc Finger Antiviral Protein. *Journal of Virology* **91** (2017).
306. Hubel, P. *et al.* A protein-interaction network of interferon-stimulated genes extends the innate immune system landscape. *Nature Immunology* **20,** 493–502 (2019).
307. Wan, Y., Shang, J., Graham, R., Baric, R. S. & Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *Journal of Virology* **94** (2020).
308. Gu, H. *et al.* Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. *Science* **369,** 1603–1607 (2020).
309. Dinnon, K. H. *et al.* A mouse-adapted model of SARS-CoV-2 to test COVID-19 countermeasures. *Nature* **586,** 560–566 (2020).
310. Muñoz-Fontela, C. *et al.* Animal models for COVID-19. *Nature* **586,** 509–515 (2020).

311. Dillard, J. A., Martinez, S. A., Dearing, J. J., Montgomery, S. A. & Baxter, V. K. Animal Models for the Study of SARS-CoV-2–Induced Respiratory Disease and Pathology. *Comparative Medicine* **73,** 72–90 (2023).
312. Rathnasinghe, R. *et al.* Comparison of transgenic and adenovirus hACE2 mouse models for SARS-CoV-2 infection. *Emerging Microbes & Infections* **9,** 2433–2445 (2020).
313. Yinda, C. K. *et al.* K18-hACE2 mice develop respiratory disease resembling severe COVID-19. *PLOS Pathogens* **17,** e1009195 (2021).
314. Chu, H., Chan, J. F.-W. & Yuen, K.-Y. Animal models in SARS-CoV-2 research. *Nature Methods* **19,** 392–394 (2022).
315. Tostanoski, L. H. *et al.* Ad26 vaccine protects against SARS-CoV-2 severe clinical disease in hamsters. *Nature Medicine* **26,** 1694–1700 (2020).
316. Imai, M. *et al.* Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. *Proceedings of the National Academy of Sciences* **117,** 16587–16595 (2020).
317. Osterrieder, N. *et al.* Age-Dependent Progression of SARS-CoV-2 Infection in Syrian Hamsters. *Viruses* **12,** 779 (2020).
318. Rockx, B. *et al.* Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model. *Science* **368,** 1012–1015 (2020).
319. Yu, P. *et al.* Age-related rhesus macaque models of COVID-19. *Animal Models and Experimental Medicine* **3,** 93–97 (2020).
320. Deng, W. *et al.* Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. *Science* **369,** 818–823 (2020).
321. Chandrashekar, A. *et al.* SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science* **369,** 812–817 (2020).
322. Vogel, A. B. *et al.* BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature* **592,** 283–289 (2021).
323. Williamson, B. N. *et al.* Clinical benefit of remdesivir in rhesus macaques infected with SARS-CoV-2. *Nature* **585,** 273–276 (2020).
324. Chen, W. *et al.* SARS-associated Coronavirus Transmitted from Human to Pig. *Emerging Infectious Diseases* **11,** 446–448 (2005).
325. Schlottau, K. *et al.* SARS-CoV-2 in fruit bats, ferrets, pigs, and chickens: an experimental transmission study. *The Lancet Microbe* **1,** e218–e225 (2020).
326. Shi, J. *et al.* Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS–coronavirus 2. *Science* **368,** 1016–1020 (2020).
327. Chu, H. *et al.* Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. *The Lancet Microbe* **1,** e14–e23 (2020).

328. Färber, I. *et al.* Investigations on SARS-CoV-2 Susceptibility of Domestic and Wild Animals Using Primary Cell Culture Models Derived from the Upper and Lower Respiratory Tract. *Viruses* **14,** 828 (2022).
329. Du, X. *et al.* Establishment of a humanized swine model for COVID-19. *Cell Discovery* **7,** 70 (2021).
330. Taylor, R. Dukes' Physiology of Domestic Animals. 12th edition, Reece WO. Cornell University Press, Ithaca, 2004, 999 pages. Price US$95.00. ISBN 8014 4238 9. *Australian Veterinary Journal* **83,** 444–444 (2005).
331. Dorward, D. A. *et al.* Tissue-Specific Immunopathology in Fatal COVID-19. *American Journal of Respiratory and Critical Care Medicine* **203,** 192–201 (2021).
332. Reuschl, A.-K. *et al.* Evolution of enhanced innate immune suppression by SARS-CoV-2 Omicron subvariants. *Nature Microbiology* **9,** 451–463 (2024).
333. Hu, B., Guo, H., Zhou, P. & Shi, Z.-L. Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology* **19,** 141–154 (2021).
334. Chan, J. F.-W. *et al.* Simulation of the Clinical and Pathological Manifestations of Coronavirus Disease 2019 (COVID-19) in a Golden Syrian Hamster Model: Implications for Disease Pathogenesis and Transmissibility. *Clinical Infectious Diseases* (2020).
335. Winkler, E. S. *et al.* SARS-CoV-2 infection of human ACE2-transgenic mice causes severe lung inflammation and impaired function. *Nature Immunology* **21,** 1327–1335 (2020).
336. Singh, D. K. *et al.* Responses to acute infection with SARS-CoV-2 in the lungs of rhesus macaques, baboons and marmosets. *Nature Microbiology* **6,** 73–86 (2020).
337. Dong, W. *et al.* The K18-Human ACE2 Transgenic Mouse Model Recapitulates Non-severe and Severe COVID-19 in Response to an Infectious Dose of the SARS-CoV-2 Virus. *Journal of Virology* **96** (2022).
338. Sia, S. F. *et al.* Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature* **583,** 834–838 (2020).
339. Chan, J. F.-W. *et al.* Surgical Mask Partition Reduces the Risk of Noncontact Transmission in a Golden Syrian Hamster Model for Coronavirus Disease 2019 (COVID-19). *Clinical Infectious Diseases* **71,** 2139–2149 (2020).
340. Barek, M. A., Aziz, M. A. & Islam, M. S. Impact of age, sex, comorbidities and clinical symptoms on the severity of COVID-19 cases: A meta-analysis with 55 studies and 10014 cases. *Heliyon* **6,** e05684 (2020).
341. Van Cleemput, J. *et al.* Organ-specific genome diversity of replication-competent SARS-CoV-2. *Nature Communications* **12,** 6612 (2021).
342. Luo, Y. *et al.* Broad Cell Tropism of SADS-CoV In Vitro Implies Its Potential Cross-Species Infection Risk. *Virologica Sinica* **36,** 559–563 (2021).

343. Nelli, R. K. *et al.* Enhanced apoptosis as a possible mechanism to self-limit SARS-CoV-2 replication in porcine primary respiratory epithelial cells in contrast to human cells. *Cell Death Discovery* **7,** 383 (2021).
344. Barjesteh, N., Abdelaziz, K. T. & Sharif, S. The role of IRF7 and NF-κB pathways in the induction of antiviral responses in chicken tracheal epithelial cells following exposure to TLR3 and 4 ligands. *The Journal of Immunology* **196,** 8–216 (2016).
345. Kim, T. H. & Zhou, H. Functional Analysis of Chicken IRF7 in Response to dsRNA Analog Poly(I:C) by Integrating Overexpression and Knockdown. *PLOS ONE* **10,** e0133450 (2015).
346. Santhakumar, D., Rubbenstroth, D., Martinez-Sobrido, L. & Munir, M. Avian Interferons and Their Antiviral Effectors. *Frontiers in Immunology* **8** (2017).