MASTER IN BIOCHEMISTRY AND MOLECULAR AND CELL BIOLOGY RESEARCH FOCUS

Impact of the APOBEC3 innate immune effectors on SARS-CoV-2 replication

Collard, Maxence

Award date:
2022

Awarding institution:
University of Namur

Faculté des Sciences

Impact of the APOBEC3 innate immune effectors on SARS-CoV-2 replication

Mémoire présenté pour l'obtention

du grade académique de master 120 en biochimie et biologie moléculaire et cellulaire

Maxence COLLARD

Janvier 2022

Université de Namur
FACULTE DES SCIENCES
Secrétariat du Département de Biologie
Rue de Bruxelles 61 - 5000 NAMUR
Téléphone: + 32(0)81.72.44.18 - Téléfax: + 32(0)81.72.44.20
E-mail: joelle.jonet@unamur.be - http://www.unamur.be

Impact des effecteurs de l'immunité innées APOBEC3 sur la réplication du SARS-CoV-2

Collard Maxence

Résumé

Les protéines APOBEC sont une famille d'enzymes effectrices de notre système immunitaire inné. Leur activité catalytique leur permet de désaminer les cytosines des acides nucléiques simples brins en uracile. Cette activité se traduit par la capacité d'introduire des mutations dans le génome des virus et des rétroéléments. Lorsque trop de mutations sont induites, cela nuit à la production de progéniture virale infectieuse. En plus de cet outil de défense lié à la désamination des cytosines, les protéines APOBEC peuvent lier différents outils impliqués dans le cycle viral et altérer la capacité des virus à se répliquer. Bien évidemment, les virus ont coévolué avec l'existence de ces mécanismes de défense et développé diverses manières de s'en protéger. La longue coexistence entre virus et désaminases a laissé des marques sur les génomes viraux et a influencé leur évolution.

Le SARS-CoV-2, le virus responsable de l'épidémie de COVID-19, a été extensivement étudié. L'ampleur de la pandémie a permis de récolter une quantité importante de données. Grâce à ces analyses, plusieurs auteurs ont observé des biais dans la distribution des nucléotides, dont une forte transition de cytosine en uracile. Cette indication nous mène à suspecter l'implication des protéines APOBEC dans l'évolution du SARS-CoV-2.

Sur base de ces observations, nous avons choisi d'étudier l'impact des protéines APOBEC sur la réplication du SARS-CoV-2. Nous avons construit différentes lignées surexprimant les différentes APOBEC et les avons infectées avec le SARS-CoV-2 afin d'y mesurer sa cinétique d'infection. Nous n'avons pu observer de différences entre les lignées surexprimant les APOBEC et les lignées ne les surexprimant pas. Nous avons aussi étudié le profil d'expression des différentes APOBEC3 dans des cellules bronchiques primaires humaines infectées afin d'identifier quelles protéines APOBEC3 pourraient être les plus pertinentes à étudier dans le cas de la restriction du SARS-CoV-2. Nous avons comparé le niveau d'expression d'ARNm et de protéines entre des cellules infectées et non infectées. Le niveau d'expression d'ARN et de protéines d'A3G est particulièrement augmenté suite à l'infection par le virus. Nous n'avons toutefois pas pu observer d'activité désaminase dans ces cellules infectées. Cela implique que le virus puisse peut-être inhiber l'activité désaminase de la protéine APOBEC3G.

Mémoire de master 120 en biochimie et biologie moléculaire et cellulaire

Janvier 2022

Promoteur : N. Gillet

Superviseur : K. Willemart

Université de Namur
FACULTE DES SCIENCES
Secrétariat du Département de Biologie
Rue de Bruxelles 61 - 5000 NAMUR
Téléphone: + 32(0)81.72.44.18 - Téléfax: + 32(0)81.72.44.20
E-mail: joelle.jonet@unamur.be - http://www.unamur.be

Impact of the APOBEC3 innate immune effectors on SARS-CoV-2 replication

Collard Maxence

Summary

APOBEC proteins are a family of effectors of our immune system. They catalyze the deamination of cytosine into uracil on single-stranded nucleic acids. This activity allows them to introduce mutations into the genome of viruses and retroelements. When a high number of mutations are induced, it impairs the production of an infectious progeny. In addition to these deamination-dependent effects, APOBEC proteins can bind to viral elements involved in their life cycle. This capacity alters the proper replication of viruses. Obviously, viruses have coevolved with defense mechanisms and developed many ways to protect themselves from the actions of the deaminases. The long coexistence between viruses and APOBEC proteins has left footprints and shaped viral evolution.

SARS-CoV-2, the virus responsible for the global outbreak of COVID-19, has been massively analyzed. Due to the magnitude of the pandemic, a huge amount of data was collected. Through these investigations, several authors have identified biases in the distribution of nucleotides. This includes a strong transition of cytosine into uracil. This observation leads us to believe that APOBEC proteins may be involved in SARS-CoV-2 restriction.

Based on these observations, we chose to study the impact of APOBEC proteins on SARS-CoV-2 replication. We constructed different cell lines overexpressing different APOBEC proteins and infected them with SARS-CoV-2 to investigate whether the APOBEC proteins could effectively limit (or promote) SARS-CoV-2 replication. We were not able to observe any difference in the SARS-CoV-2 replication kinetics between cell lines overexpressing APOBEC proteins and those not overexpressing them. In order to guide us on which APOBEC3 would be most relevant to study in the context of SARS-CoV-2 replication, we decided to study their expression profile in primary human bronchial cells infected with SARS-CoV-2. We compared infected cells with uninfected cells and it turned out that the expression of APOBEC3G is particularly upregulated at the mRNA and protein levels. We could not observe any increased deamination activity in the infected cells despite an increase of APOBEC3G protein level. This raises the possibility that the virus can inhibit the deaminase activity of the APOBEC3G protein.

Master 120 thesis in biochemistry and molecular and cellular biology

January 2022

Promoter: N. Gillet

Supervisor: K. Willemart

Acknowledgment

J'aimerais remercier toutes les personnes qui ont contribuées de près ou de loin au succès de mon mémoire mais aussi à mes études en général.

Je tiens particulièrement à remercier le Pr. Nicolas Gillet de m'avoir accueilli dans son laboratoire, aussi bien lors de mon premier bénévolat en URVI que lors de ces dix mois de mémoire. Je te remercie pour ta confiance, nos discussions et tes corrections. Sans lui, je ne serais peut-être pas ici pour écrire ces quelques mots de remerciements.

Je remercie également Kévin, mon encadrant, pour toutes ces heures à m'apprendre et m'expliquer le travail en laboratoire. Je te remercie pour toutes tes réponses à ces questions parfois un peu bêtes, pour toutes ces heures non-comptées à m'encadrer. Tu as toujours été disponible pour moi, peu importe le moment de la journée ou de la semaine. Merci d'avoir été à mon écoute, même dans mes mauvaises phases. Merci, pour cet encadrement, pour un premier, tu as été top !

Je te remercie aussi Noémie. Finalement, tu auras été ma première encadrante. Tu m'as appris les toutes premières bases du travail en laboratoire lors de mon premier stage. Même maintenant, je sais que je peux toujours compter sur toi pour répondre à n'importe lesquelles de mes questions.

Je remercie tous les membres de l'URVI pour l'encadrement qu'ils m'ont apportés. Merci Laeti, merci Hélène pour ces petits coups de pouces et tous ces petits renseignements. Merci Damien, pour tous tes précieux conseils. Merci Alexis, d'être comme tu es, ne change pas. Merci Astrid, Benoit, Nicolas, Laura pour votre bonne humeur. Merci à tous de m'avoir intégré dans l'équipe. Merci à toute l'équipe APOBEC, Alex, Florian, Sarah, Zoé, on peut toujours compter sur vous.

Merci à mon acolyte de toujours, Boggy, tu auras été le meilleur support pour moi durant ce master.

Merci à ma belle-famille pour toute l'attention et le support qu'elle m'apporte. Merci Alix de m'accompagner chaque jour.

Enfin, tout particulièrement merci à ma maman. Tu as toujours été là pour moi. Tu as sacrifié bien plus que tu ne l'aurais dû pour mon confort. Merci d'avoir toujours cru en moi, de m'avoir soutenu tout au long de mes études mais aussi de ma vie.

Merci à toutes les personnes qui ont fait de moi l'homme que je suis aujourd'hui.

List of abbreviations

%	Percent
°C	Degree Celsius
µ	Micro
A	Ampere
A	Adenosine
AAV	Adeno-Associated Virus
ACE2	Angiotensin-Converting Enzyme 2
ADAR	Adenosine Deaminase Acting On RNA
AID	Activation-Induced Domains
APE	Apurinic-Apyrimidinic Endonucleases
ApoB	Apolipoprotein B
APOBEC	Apolipoprotein B mRNA Editing Enzyme Catalytic Polypeptide-Like
ATP	Adenosine Triphosphate
BCA	Bicinchoninic Acid
BHQ1	Black Hole Quencher-1
BIV	Bovine Immunodeficiency Virus
BPE	Bovine Pituitary Extract
BSA	Bovine Serum Albumin
C	Cytosine
CAEV	Caprine Arthritis Encephalitis Virus
CBFβ	Core-Binding Factor
CDK4	Mouse Cyclin-Dependent Kinase
cDNA	Complementary DNA
CDS	Coding DNA Sequence
CM	Convoluted Membrane
COVID-19	Coronavirus Disease 19
Ct	Cycle Threshold
Cy5	Cyanine 5
DMEM	Dulbecco's Modified Eagle Medium
DMS	Double Membrane Open Spherules
DMV	Double Membrane Vesicle
DNA	Desoxyribonucleic Acid
DPI	Days Post Infection
ds	Double-Stranded
DTT	Dithiothreitol
E. coli	Escherichia Coli
EBV	The Epstein-Barr Virus
EDTA	Ethylenediaminetetraacetic Acid
eEF1α	Eukaryotic Elongation Factor 1 Alpha
EIAV	Equine Infectious Anemia Virus
EMEM	Eagle's Minimal Essential Medium
ER	Endoplasmic Reticulum
ERGIC	Endoplasmic Reticulum – Golgi Intermediate Compartment
EV-A71	Enterovirus A71
FACS	Fluorescence-Activated Cell Sorting
FAM	6-Carboxyfluorescein

FBS	Fetal Bovine Serum
FCS	Fetal Calf Serum
FIV	Feline Immunodeficiency Virus
Fw	Forward
G	Guanosine
g	Gravitational Force Equivalent
GAP	GTPase Activating Protein
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GFP	Green Fluorescent Protein
GTP	Guanosine Triphosphate
GTPase	Guanosine Triphosphate Hydrolase
h	Hour
HA-tag	Human Influenza Hemagglutinin Tag
HBEC3-KT	Human Bronchial Epithelial Cell
HCoV	Human Coronavirus
HED	HEPES-EDTA-DTT
HEK 293T	Human Embryonic Kidney 293T Cells
HEPES	4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid
HIV	Human Immunodeficiency Virus
HMM	High Molecular Mass
HPRT	Hypoxanthine Phosphoribosyltransferase
HPV	Human Papilloma Virus
HRP	Horseradish Peroxidase
hRSV	Human Respiratory Syncytial Virus
Hsp90	Heat Shock Protein
HSV1	Herpes Simplex Virus 1
HTLV-1	Human T-Lymphotropic Virus 1
IDL	Intermediate Density Lipoprotein
IFN	Interferon
IL	Interleukin
kDa	Kilodalton
KSHV	Kaposi’s Sarcoma Herpesvirus
L	Liter
LDL	Low-Density Lipoprotein
LINE	Long Interspersed Nuclear Element
LMM	Low Molecular Mass
LTR	Long Terminal Repeat
M	Mole
m	Milli
MERS-CoV	Middle East Respiratory Syndrome Coronavirus
min	Minute
MLV	Murine Leukemia Virus
MMTV	Mouse Mammary Tumor Virus
MOI	Multiplicity Of Infection
mRNA	Messenger RNA
MuV	Mumps Virus
MV	Measles Virus
MVM	Minute Virus of Mice
MVV	Maedi-Visna Virus
NaOH	Sodium Hydroxide

NC	Nucleocapsid
NF1	Neurofibromin
Nsp	Nonstructural Protein
ORF	Open Reading Frame
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
pH	Potential Of Hydrogen
PI	Post-Infection
pmol	Picomole
PMSF	Phenylmethylsulfonyl Fluoride
PRF	Programmed Ribosomal Frameshift
RBx2	RING-Box Subunit 2
RdRp	RNA-Dependent RNA Polymerase
RIG-1	Retinoic Acid-Inducible Gene-1
RIPA	Radioimmunoprecipitation Assay
RNA	Ribonucleic Acid
rpm	Round Per Minute
RSV	Rous Sarcoma Virus
RTC	Replication/transcription Complex
RTqPCR	Reverse Transcriptase Quantitative PCR
Rv	Reverse
S	Svedberg
SARS-CoV	Severe Acute Respiratory Syndrome Coronavirus
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SCR	Scrambled
sec	Second
sgRNA	Subgenomic RNA
Sh	Short Hairpin
SINE	Short Interspersed Nuclear Element
SIV	Simian Immunodeficiency Virus
SNP	Single-Nucleotide Polymorphism
TBP	TATA-Binding Protein
TBS-T	Tri-Buffered Saline - Tween
TCID50	50% Tissue Culture Infectious Dose
TMPRSS2	Transmembrane Protease Serine 2
TNF-α	Tumor Necrosis Factor A
tRNA	Transfer RNA
TRS-B	Body Transcription Regulatory Sequence
TRS-L	Leader Transcription Regulatory Sequence
U	Uracil
UDG/UNG2	Uracil-DNA Glycosylase
UTR	Untranslated Region
V	Volt
Vif	Virion Infectivity Factor
VLDL	Very Low-Density Lipoprotein
W	Watt
WB	Western Blot
WT	Wild Type
ZDD	Zinc-Dependent Deaminase Domain
$ZnCl_2$	Zinc Chloride

Table of figures

Figure S14: Map of the lentiviral plasmid bearing the scrambled shRNA transgene and blasticidin resistance gene.

Figure S15: Map of the lentiviral plasmid bearing the scrambled shRNA A3G transgene and blasticidin resistance gene.

Figure S16: Map of the lentiviral plasmid bearing the shRNA cloning site and puromycin resistance gene.

Figure S17: Schematic representation of the genome of the SARS-CoV-2 strain used for Vero E6 and HBEC3-KT-ACE2 infections. The strain was isolated from a Belgian patient infected with SARS-CoV-2 in March 2020.

Figure S18: RNA quantification of A3 gene by RTqPCR upon HBEC3-KT-ACE2 infection by SARS-CoV-2 relative to housekeeping genes (HPRT/GAPDH/TBP) mRNA levels in replicate n°1.

Figure S19: RNA quantification of A3 gene by RTqPCR upon HBEC3-KT-ACE2 infection by SARS-CoV-2 relative to housekeeping genes (HPRT/GAPDH/TBP) mRNA levels in replicate n°2.

Figure S20: RNA quantification of A3 gene by RTqPCR upon HBEC3-KT-ACE2 infection by SARS-CoV-2 relative to housekeeping genes (HPRT/GAPDH/TBP) mRNA levels in replicate n°3.

Figure S21: Assessment of A3G upon HBEC3-KT-ACE2 infection by SARS-CoV-2 in replicate n°2.

Figure S22: Assessment of A3G upon HBEC3-KT-ACE2 infection by SARS-CoV-2 in replicate n°3.

Figure S23: Assessment of A3G deaminase activity upon HBEC3-KT-ACE2 infection by SARS-CoV-2 in replicate n°1.

Figure S24: Assessment of A3G deaminase activity upon HBEC3-KT-ACE2 infection by SARS-CoV-2 in replicate n°3.

List of tables

Table of contents

Introduction

APOBEC proteins, or Apolipoprotein B mRNA editing enzyme catalytic polypeptide-like, are Zn^{2+} dependent cytidine deaminases. These proteins catalyze the deamination of cytosine into uridine (C-to-U) of single-stranded nucleic acids (Figure 1) (Salter *et al.*, 2016). APOBEC proteins are actually a family of proteins made of eleven members, divided into five subfamilies: activation-induced deaminase (AID), APOBEC1 (A1), APOBEC2 (A2), seven APOBEC3 (A3A-A3H) and APOBEC4 (A4). While historically discovered in the context of the edition of the Apolipoprotein B mRNA, most of these proteins are effectors of the innate immune system. Indeed, deamination is one of the many tools used by the innate immune system to fight viruses. By mutating viral genetic material, APOBEC proteins are able to restrict virus but also to act against endogenous retroelements. These proteins are therefore major effectors of the innate immune system of eukaryotic organisms (Salter *et al.*, 2016). Due to their capacity to deaminate cytosine into uridine, APOBEC proteins can create biases in codon distribution. These signatures are called the APOBEC evolutionary footprint and are found in many viruses (Poulain *et al.*, 2020).

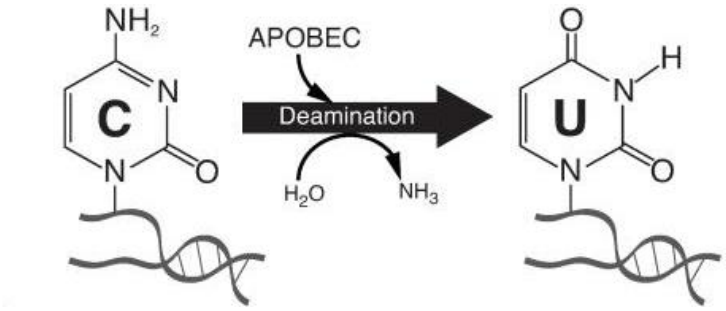

Figure 1: APOBEC proteins deaminate cytosine into uridine on single-stranded nucleic acids (Swanton *et al.*, 2015).

APOBEC family discovery

In 1987, the study of two related proteins, namely ApoB100 and ApoB48, led to the discovery of a new mechanism of mRNA editing. The protein ApoB48 weighs 48kDa and is a shorted version of the ApoB100. The ApoB48 is produced by enterocytes whereas the ApoB100 is produced by liver cells. Importantly, these two proteins were shown to be encoded by the same gene, the Apolipoprotein B (ApoB) gene (S. H. Chen *et al.*, 1987; Powell *et al.*, 1987). The difference between the two is due to a post-transcriptional modification of the ApoB mRNA at position 6666. When unedited, the ApoB mRNA leads to the production of the ApoB100 protein. But in enterocytes, the cytosine at position 6666 can be deaminated into an uracil leading to the generation of a premature stop codon. The glutamate codon, CAA, is converted into UAA leading to the production of a truncated protein as UAA is read as a stop codon (Figure 2). This truncated protein is the ApoB48 protein and has a different biological function. The enzyme responsible for this targeted deamination was isolated in 1993 and named APOBEC1 in 1995 (Davidson *et al.*, 1995; Teng *et al.*, 1993). This finding led to the identification of multiple related genes and the expansion of the APOBEC family to the eleven members we know in humans.

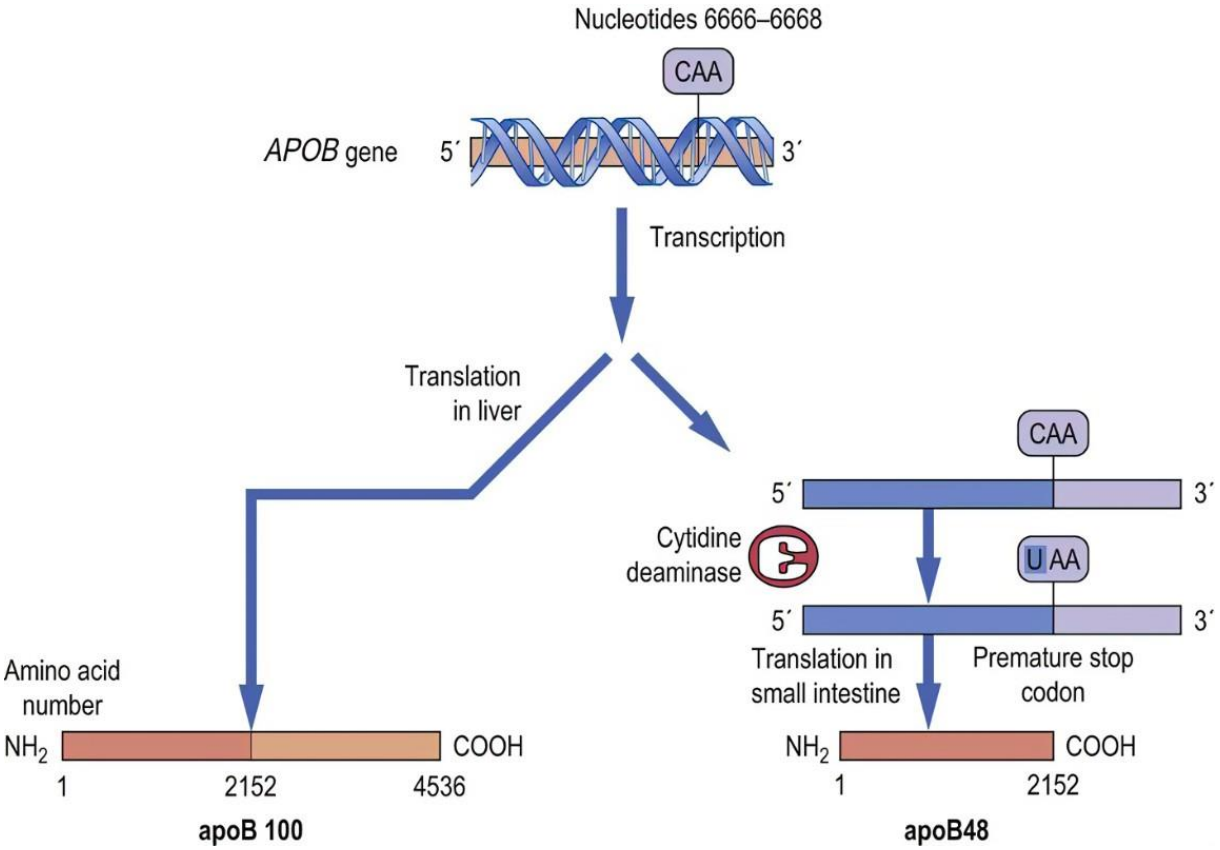

Figure 2: Post-transcriptional editing of ApoB mRNA in human enterocytes by cytidine deaminase leads to a premature codon and allows tissue-specific expression of ApoB48 transcript (Patton *et al.*, n.d.).

Origin and evolution of the APOBEC family

APOBEC proteins are conserved and widespread in the animal kingdom. The family is related to deaminases found in bacteria, yeasts and plants and shares the same conserved domains for zinc coordination (Conticello *et al.,* 2005). The AID gene is believed to be the first ancestral member of the APOBEC family that emerged. Indeed, it was already expressed 500 million years ago in jawless fish's lymphocytes. (Rogozin *et al.*, 2007). From the AID gene, duplication events led to the diversification of the APOBEC family in vertebrates. A2 emerged from it in bony fishes but it also allowed the diversification of AID functions (Conticello *et al.,* 2005). For example, AID homologs are found in sea urchins and brachiopods. In this case, they have more of an antimicrobial function whereas the homologs found in vertebrate lineages are involved in antibody diversification (Liu *et al.,* 2018). Homologs of our A2 are found in birds, amphibians and ray-finned fish (Conticello *et al.,* 2005; Hsu, 2016). Others AID duplications events led to A1 and A3 rise (Conticello *et al.,* 2005). A1 homologs are found in birds, reptiles, amphibians and lungfishes (Krishnan *et al.,* 2018). The A3 family is unique to placental mammals and is therefore absent from marsupials and monotremes (Münk *et al.,* 2012). The place of A4 is particular and is actually not well understood. Indeed, recent bioinformatical analyses have identified A4 homologs in jellyfishes and some algae but not in fishes (Krishnan *et al.,* 2018). It is however present in frogs, birds and mammals. Curiously, A4 seems to be phylogenetically closer to A1. There are still many unknowns around A4 and the mystery of its origin has yet to be solved. Hence, these recent analyses suggest a possibly older origin of the APOBEC protein family (Figure 3) (Rogozin *et al.,* 2005).

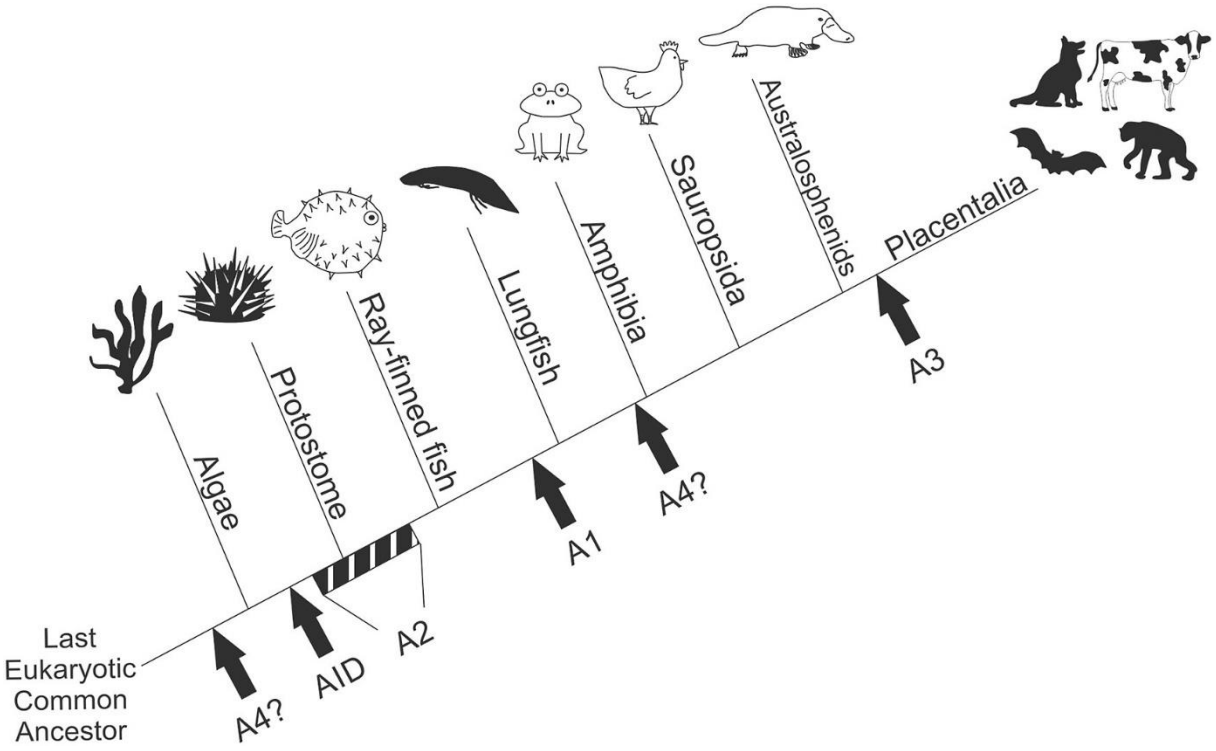

Figure 3: Evolutionary origin of the APOBEC family. A4 origin is uncertain. A2 emergence may have occurred before or after protostome-deuterostome separation (Ratcliff & Simmonds, 2021).

Three A3 paralogs are found in placental mammals and have undergone multiple diversification events such as duplication, fusion and loss. These differential evolution processes led to the diversity of paralogues and would be at the origin of the seven APOBEC3 (A3) found in humans (Figure 4) (Krishnan *et al.*, 2018; LaRue *et al.*, 2009; Münk *et al.*, 2012). The number of A3 in mammals is variable among mammals: only one A3 is present in rodents while 18 are found in bats (Hayward *et al.*, 2018). Two deaminase domains are unique to A3 and certainly originate from duplication events followed by gene fusions (Münk *et al.*, 2012). A3 genes evolve at a fast rate, a common feature for effectors involved in antiviral responses (Sawyer *et al.*, 2004).

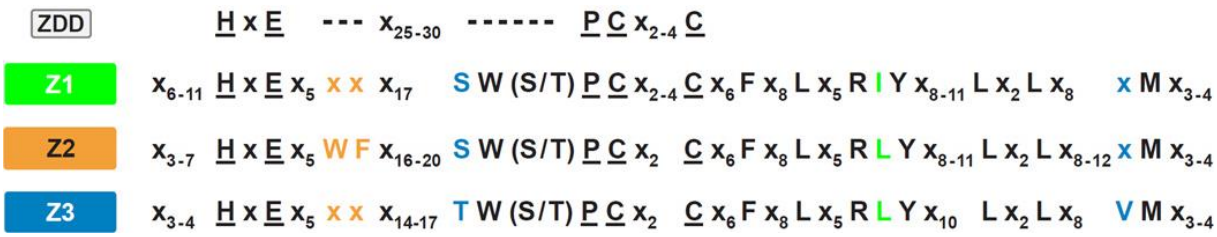

Figure 4: Zinc-dependent deaminase domain (ZDD). Canonical ZDD is represented in white. A3 paralogs certainly emerged from duplication events and are depicted in colors. Z1 and Z2 are shared in the A3 subfamily. Z3 is unique to A3H (Salter *et al.*, 2016).

Structure conservation and diversity

Although members of the same family, APOBEC proteins show differences and similarities. Each member shares one or two highly conserved Zn^{2+} dependent deaminase domains (ZDD).

Two domains are present in A3B, A3DE, A3F and A3G while a single ZDD is present in the other members of the APOBEC family (Figure 5). This deaminase domain is implicated in virus restriction but also in various biological functions such as lipid metabolism and antibody diversification. A highly conserved His-X-Glu-X_{23-28}-Pro-Cys-X_{2-4}-Cys motif forms the catalytic core of the enzyme and coordinates zinc ions (Figure 4) (Salter *et al.*, 2016). The glutamate residue deprotonates a water molecule. This water molecule serves as a proton and hydroxyl donor. The carbon 4 of the cytidine pyrimidine ring undergoes a nucleophilic attack from the zinc-stabilized hydroxide ion. The amine group is replaced by a carbonyl group resulting in the deamination of the cytidine into uridine (Figure 6) (Harris & Dudley, 2015).

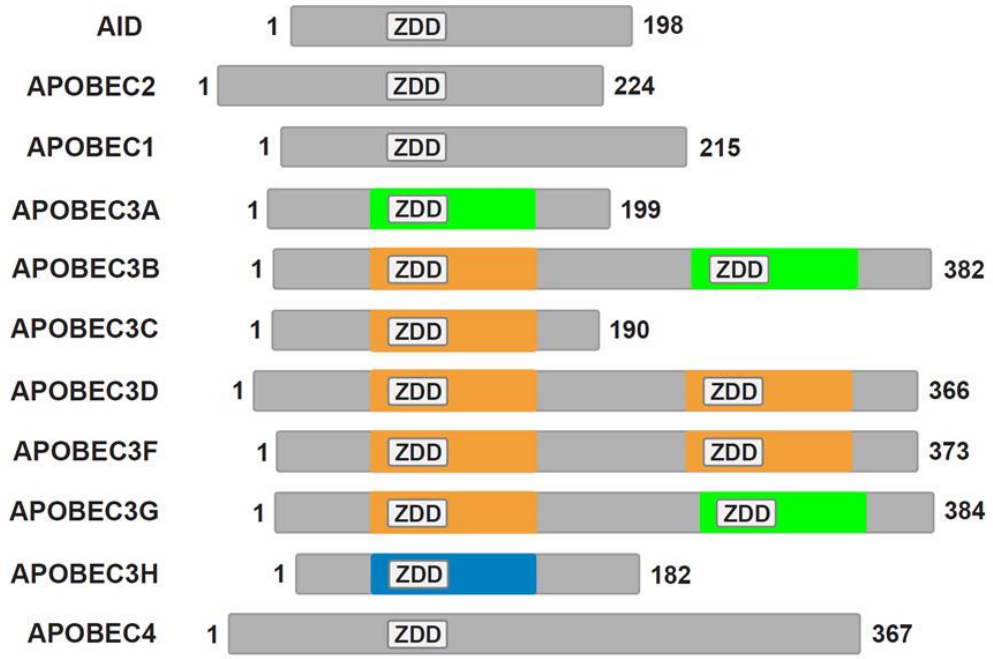

Figure 5: Eleven APOBEC members are expressed in humans. AID, A2, A1, A3A, A3C,A3H, A4 bears a unique ZDD. A3B, A3DE, A3F and A3G bears two ZDD. APOBEC proteins are represented to scale and amino acid size is indicated on the c-terminus region. (Salter *et al.*, 2016).

The deaminase domain is found within a super-secondary structure made of five β-strands and six α-helices and shows a globular architecture. The β-strands make a hydrophobic β-sheet core while the α-helices surround the core. In the presence of two deaminase domains, a flexible linker joins them. Not all domains are catalytically active but still have functions. N-terminus ZDD of A3B, A3F and A3G are catalytically inactive. These domains are however necessary for nucleic acids binding. Patches of positively charged and aromatic residues allow stabilization and interaction with the nucleic acid backbone, which is negatively charged. Moreover, the C-terminus domain loses its activity in the absence of the N-terminus domain. Minor variations in the residues allow a variety of lengths, compositions and spatial locations of the secondary structure. For example, the β-strand 2 can be split. The interruption of the β-strand allows the presence of a short bulging loop that will modify interactions. This is the case in both A3A deaminase domains and in the C-terminus deaminase domains of A3B and A3G. This slight diversity enables the distinction between the five APOBEC subfamilies and is certainly the reason that leads to the various functions, regulations or substrate selections (Harris & Dudley, 2015; Salter *et al.*, 2016). Unfortunately, the low diversity makes the production of specific antibodies difficult, especially for the A3 subfamily (Burns *et al.*, 2015).

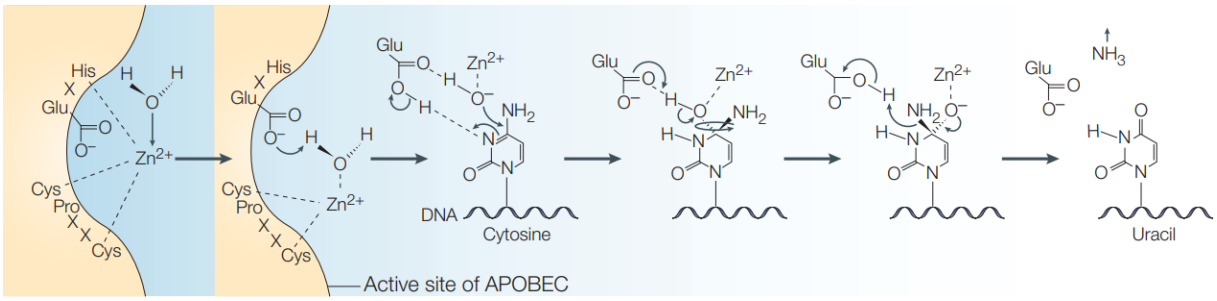

Figure 6: Suggested mechanism for APOBEC deamination based on yeasts and bacteria deaminase mechanisms. Cysteine and histidine coordinate zinc ions. A water molecule is deprotonated by the glutamate residue and serves as a proton and hydroxyl donor. Protonation of N3 destabilizes the double bond with C4 which undergoes nucleophilic attack from the hydroxide ion. The amino group is protonated and the C-N bond is cleaved leading to the formation of a double bound on C4 and the release of ammonia (Harris & Liddament, 2004).

Substrate selection and deamination hotspots

APOBEC family members have different binding activities and preferences in terms of substrate selection. Indeed, all APOBEC proteins have the capacity to bind single-stranded DNA (ssDNA) and single-stranded RNA (ssRNA) but only A1, AID, A3A and A3G are able to deaminate ssRNA. The other ones are only able to deaminate ssDNA (Salter *et al.*, 2016; Sharma *et al.*, 2015). In any case, a minimum of five continuous nucleotides are needed for APOBEC activity (Harris & Dudley, 2015).

Besides their specificities to DNA or RNA, APOBEC proteins have also a favorite nucleotidic sequence to focus on. These target contexts are called motifs and are hotspots for deamination processes. Most APOBEC proteins prefer to target cytosine preceded by a thymine (5'-TC-3'), A3G prefers to target cytosine preceded by another cytosine (5-CC-3') and AID prefers to target cytosine preceded by a purine, itself preceded by an adenosine or a thymine (5'-WRC-3') (Pham *et al.*, 2003; Salter *et al.*, 2016). A1 is particular as it needs a cofactor for its activity. The A1 complementation factor (A1CF) binds an eleven-nucleotides sequence (the mooring sequence) four to six nucleotides downstream of the target. Then, A1 deaminates in the context of a cytosine preceded by an adenosine (5'-AC-3') (Salter *et al.*, 2016; Smith *et al.*, 2012). Although the base identity preceding the cytosine (minus-one site) is a key factor, the nucleobases at position minus-two and plus-one sites are also determinants for the target selection. For example, A3A, A3F and A3H would preferentially target 5'-TC-3' flanked by an adenosine or a thymine over a guanosine. Moreover, the DNA integrity and the secondary structure of the target are influential in substrate selection. Indeed, most APOBEC proteins will prefer to deaminate a 5'-TC-3' target in ssDNA regions over a motif found in dsDNA such as loop or bulge regions (Holtz *et al.*, 2013; McDaniel *et al.*, 2020).

In addition to different interactants, either cellular or viral, these differences will be involved in determining the functions of APOBEC proteins.

Regulation of APOBEC proteins

The APOBEC family members may share common regulatory mechanisms as well as being very different. But in any case, APOBEC proteins activity has to be tightly regulated. In fact, due to their capacity to bind nucleic acids, uncontrolled activity of APOBEC proteins could leads to mutations in self DNA and could jeopardize genomic integrity (Salter *et al.*, 2016).

These regulatory mechanisms can be brought by different processes such as molecular interactions and bindings but it also involves the subcellular localization (Salter *et al.,* 2016; Smith *et al.,* 2012).

APOBEC proteins are differently distributed within the different cellular compartments. Some are cell-wide distributed while some are restricted to a particular subcellular compartment. A3B is the only APOBEC protein limited to the nucleus. A3DE, A3F and A3G are cytoplasmic and the other ones are both nuclear and cytoplasmic. This distribution is made possible thanks to different signal mechanisms of nucleo-cytoplasmic trafficking (Salter *et al.,* 2016; Smith *et al.,* 2012).

For instance, A1 and AID, two APOBEC proteins that have to be in the cell nucleus for their activity, show regulatory localization signals. Both A1 and AID possess an N-terminal nuclear localization signal and a C-terminal cytoplasmic retention signal that retains these proteins in the cytoplasm (Bennett *et al.,* 2006; Ito *et al.,* 2004). Besides these forms of signals, there are other mechanisms that are involved in subcellular localization such as the A1 cofactor, A1CF. It is done through direct or indirect mechanisms and this cofactor has a lot of roles in A1 regulation aside from its localization. Furthermore, AC1F itself is regulated through many different mechanisms like its phosphorylation state or its own subcellular localization, also dictated by a nuclear localization signal. This adds another level of regulation on A1 activity. Subcellular localization involves a lot of interactants such as the heat shock protein 90 (Hsp90) or the eukaryotic elongation factor 1 α (eEF1α) in the AID case. These molecules will help in retaining APOBEC proteins in the cytoplasm. AID and A1 will therefore only be imported when necessary by import mechanisms such as importin (Hogg, 2010; Salter *et al.,* 2016; Smith *et al.,* 2012).

On the opposite, A3B is kept in the nucleus. This is set through a classical localization signal, but also thanks to two distinct nuclear localization sequences on the N-terminus domain. This import system is particular as it differs from the classical motif that is typically positively charged and made of basic amino acids. In this case, the two import surfaces do not contain any positively-charged residues (Lackey *et al.,* 2012; Salamango *et al.,* 2018).

Besides that, there are proteins that do not have access to the nucleus. For example, A3G is strictly restricted to the cytoplasm. Even during the mitosis, when the nucleus is dismantled, A3G never meets chromosomal DNA thanks to a strong retention signal found at amino acid positions 113 to 128 of the protein (Bennett *et al.,* 2008; Smith *et al.,* 2012).

In addition to the subcellular localization, APOBEC expression is differentially expressed in the tissues and the level of expression is highly cell type-specific (Meshcheryakova *et al.,* 2021)

Among the many molecular interactions that can regulate APOBEC activity, we can find RNA binding. Indeed, APOBEC proteins can bind ssRNA. This interaction allows the cell to lock the deaminase activity. However, the way in which RNA regulates APOBEC activity is not yet fully understood. Nevertheless, several mechanisms are hypothesized. First, a competitive mechanism is suggested. The deaminase domains could be sterically cluttered by the RNA, which would prevent access to the catalytic site for the targets. Second, although some deaminase domains are catalytically dead, they are still able to bind nucleic acids. This association could induce conformational changes that would allosterically regulate the APOBEC activity. The conformational changes could directly displace DNA from the catalytic site or alter the site in such a way that the DNA target can no longer lodge in it (Figure 7). This model is therefore only applicable to APOBEC proteins bearing two deaminase domains (McDougall & Smith, 2011; Smith *et al.,* 2012). Third, in connection with the localization-based regulation, some APOBEC such as A3B, A3DE, A3F, A3G and A3H can show two

different forms in the cytoplasm: a low molecular mass complex (LMM) and a high molecular mass complex (HMM). In the second one, RNA is bound to APOBEC proteins and form this heavier complex. In this second form, RNA-APOBEC complexes are catalytic-inactive and sequestered in P-bodies and stress granules, in the cytoplasm (Gallois-Montbrun *et al.*, 2007; Smith *et al.*, 2012; Wichroski *et al.*, 2006). The formation of such complexes is cytokine stimulation-dependent. Indeed, interleukin-2 (IL-2), interleukin-7 (IL-7) and interleukin-15 (IL-15) stimulate the formation of HMM while Tumor necrosis factor α (TNF- α) and Polyinosinic:polycytidylic acid, an immunostimulant, promoted their dissociation to LMM, the catalytically active stance of the complex (Kreisberg *et al.*, 2006; Stopak *et al.*, 2006). The presence of APOBEC proteins in P-bodies could point to other functions of APOBEC proteins. Indeed, A3G has been shown to promote the dissociation of miRNA from mRNA and to interact with proteins from the RNA interference pathway (Gallois-Montbrun *et al.*, 2007; Huang *et al.*, 2007).

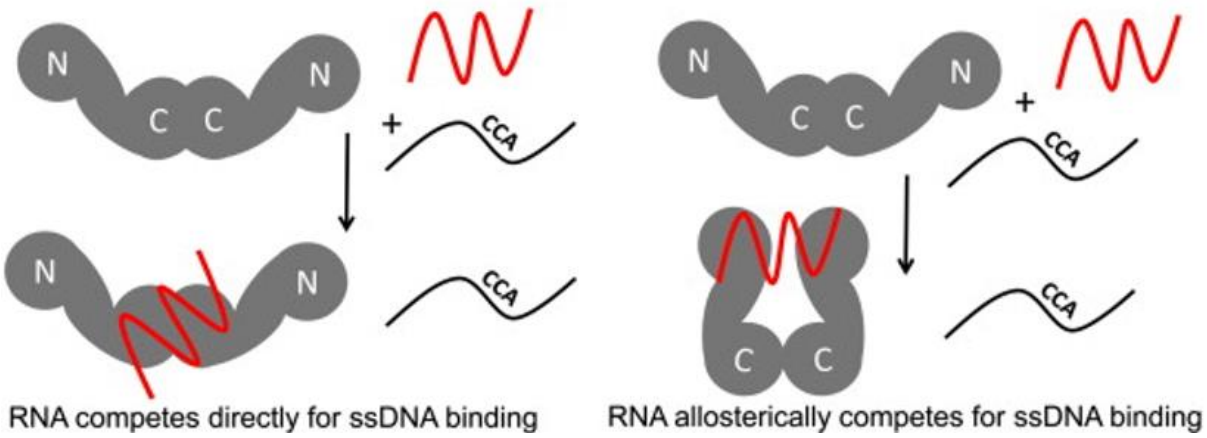

Figure 7: RNA regulates APOBEC dimers activity either by direct competition or by allosterically competing with substrates. APOBEC monomers are represented with N- and C-terminus regions. RNA are depicted in red, APOBEC substrates are depicted in black (Smith *et al.*, 2012).

In addition, A3 proteins such as A3A or A3G are well known to be dependent on interferon (IFN) induction. IFN are major modulators of A3 proteins and are produced upon detection of viral pathogens (Covino *et al.*, 2018; Greenwell-Wild *et al.*, 2009). There are many other stimuli that can regulate APOBEC activity and expression, either directly or through different pathways such as the IFN pathway. For example, 5'ppp-dsRNA recognition by RIG-1 receptors can mediate A3 expression. Nevertheless, the molecular signaling pathways underlying the expression of APOBEC proteins and regulatory mechanisms of its activity are still poorly understood (Covino *et al.*, 2018).

Finally, the oligomeric status of APOBEC proteins could be a key regulator in APOBEC activity, cellular distribution and their capacity to bind nucleic acids. Some APOBEC members, namely A3B, A3DE, A3F, A3G and A3H are capable to form homodimers that will modify APOBEC behavior (J. Li *et al.*, 2014). These oligomerization events can take place directly through protein-protein interactions or indirectly through nucleic-acid bridged interaction. The oligomeric status is actually correlated to APOBEC proteins sequestration in P-bodies and to the capacity to be packaged in HIV-1 virions (Gallois-Montbrun *et al.*, 2007; Huthoff *et al.*, 2009; Wichroski *et al.*, 2006). It is suggested that this status could modify the orientation of loops in the catalytic site of APOBEC proteins. A1 forms heterodimers with its cofactor, A1CF, that are necessary for the editing of the ApoB mRNA (Jarmuz *et al.*, 2002; Salter *et al.*, 2016).

The actions of APOBEC proteins must be permanently monitored by the cell as uncontrolled activity of APOBEC proteins or off-target effects can be genotoxic. Misregulation of APOBEC

proteins is indeed associated with malignant transformation and oncogenesis. However, the relation between APOBEC proteins and cancer is not fully understood. The link has yet to be elucidated (Henderson & Fenton, 2015; Smith *et al.,* 2012; Yamanaka *et al.*, 1995, 1997).

Members of the APOBEC family

AID, or Activation-induced deaminase, human gene is located on chromosome 12 and encodes a 198 amino acid long protein (Muto *et al.,* 2000). It is mostly expressed in lymphocytes. Contrary to most other members of the APOBEC family, AID functions do not directly target viruses. AID is an essential component of the adaptive immune system and is involved in somatic hypermutation, class switch recombination and gene conversion. In fact, the adaptive immune system is able to produce a large diversity of immunoglobulins. AID contributes to the diversification of antibodies by punctually mutating cytidines into uridines in immunoglobulin genes. The conversion leads to the recruitment of DNA repair enzymes. These DNA repair mechanisms create diversity by low-fidelity excision and mismatch repair, expanding the range of antibody possibilities (Arakawa *et al.,* 2002; Fugmann & Schatz, 2002; Papavasiliou & Schatz, 2002). AID dysregulation leads to genomic instability and is involved in malignant transformations. Moreover, by being involved in class switch recombination, defects in AID can lead to hyper-IgM syndrome type 2. In that case, immunoglobulins are unable to recombine constant region with variable region of the immunoglobulin gene leading to the accumulation of IgM and immune deficiency (Revy *et al.,* 2000).

APOBEC1 is the first discovered member of the family (Powell *et al.,* 1987). The A1 gene is also found on chromosome 12 and encodes a 236 amino acid long protein of 27kDa. In humans, A1 is only expressed in the small intestine (Teng *et al.,* 1993). A1 expression is also found in the liver of rodents, dogs and horses (Greeve *et al.,* 1993). A1 is primarily known to be involved in lipid biology. Indeed, A1 is able to posttranscriptionally edit the apolipoprotein B mRNA. Two distinct forms are found, the unedited mRNA will be translated into ApoB100 while the edited mRNA (deamination at position 6666) will be translated into ApoB48. These two proteins are involved in lipid metabolism and regulate the level of lipoproteins in the plasma. ApoB100 allows the production of very low-density lipoprotein particles (VLDL), which will then be metabolized into intermediate density lipoproteins (IDL) and low-density lipoproteins (LDL). LDL is also known as the "bad" form of cholesterol and is associated with an increased risk of atherosclerosis. On the contrary, lipoprotein particles associated with ApoB48 proteins will not be metabolized into LDL. mRNA editing takes place immediately after transcription, in the nucleus, and involves the transition from a glutamine codon, CAA, into a premature stop codon, UAA. The ApoB mRNA is not the only substrate of A1 (Davidson & Shelness, 2000; Salter *et al.,* 2016). The cytosine on position 2914 of the neurofibromin (NF1) mRNA is also a target for A1 deamination. It leads to the transition of an arginine amino acid (CGA) into a stop codon (UGA). Standard NF1 is a tumor suppressor encoding a GTPase activating protein (GAP). The truncated NF1 loses its GAP domain and is no longer a tumor suppressor. A higher level of truncated NF1 proteins is observed in solid tumors of neurological origin. The biological significance of this editing event is unknown, especially since A1 expression is thought to be restricted to the small intestine (Mukhopadhyay *et al.,* 2002; Skuse *et al.,* 1996; Smith *et al.,* 2012).

As already explained, these two editing events are dependent on the mooring sequence found on mRNA substrates. It allows the docking of the A1 cofactor, A1CF (Figure 8). Indeed, the edition of these two mRNA needs both the catalytic activity of A1 and AC1F. In addition, AC1F is needed for A1 nuclear localization. In contrary to A1, AC1F expression is expressed in many

different tissues, even those not expressing A1. This may imply other unknown biological functions for AC1F (Salter *et al.*, 2016). In addition, AC1F loss is lethal while that of A1 is not (Blanc *et al.*, 2005; Hirano *et al.*, 1996). A1 activity activation can also be brought by the RNA binding motif protein 47 (RBM47) (Fossat *et al.*, 2014).

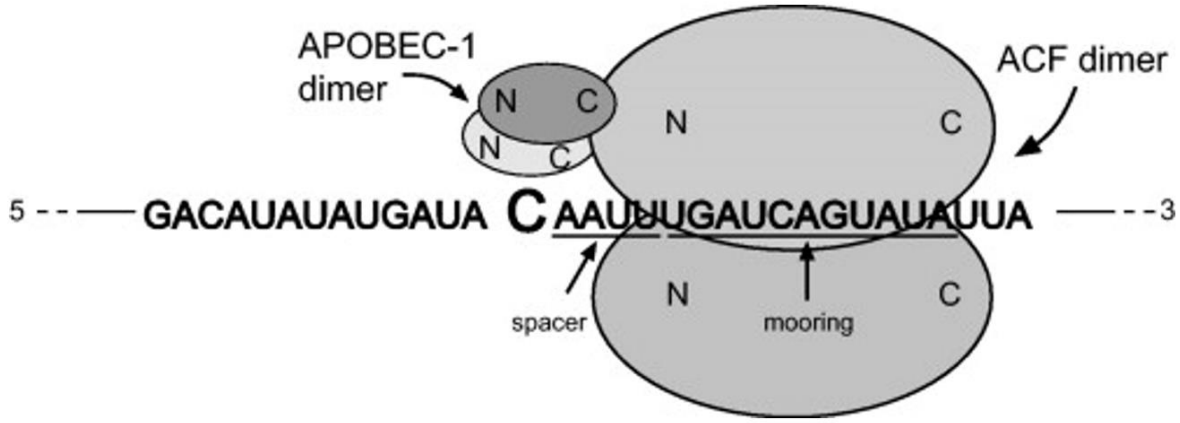

Figure 8: APOBEC1 edition requires binding to A1CF and its docking on the mooring sequence (Smith *et al.*, 2012).

APOBEC2 is one of the less-studied members of the APOBEC family. It is however the first and only member whose crystal structure has been completely elucidated, even before a function was assigned to it (Prochnow *et al.*, 2006). A2 gene is located on chromosome 6 and is 224 amino acids long (Smith *et al.*, 2012). It is only expressed in skeletal muscles and cardiac tissue (Liao *et al.*, 1999). Although used as a model to study the structure of the members of the family, its functions seem to be quite different from those of other APOBEC proteins. In fact, A2 shows a low cytidine deaminase activity (Smith *et al.*, 2012). A2 may be involved in the proper differentiation and development of muscles. Its biological functions remain uncertain. A2 deficiency is associated with muscle defects and lower body mass both in zebrafish and mice (Liao *et al.*, 1999; Sato *et al.*, 2010). It has also been observed that A2 deficiency can cause mitochondrial defects and leads to increased mitophagy in muscle cells. A2 could therefore be involved in mitochondrial homeostasis (Sato *et al.*, 2018).

APOBEC4 gene is found on chromosome 1 and encodes a 367 amino acids long protein. Very little is known about A4. During AID homology studies, bioinformatics actually discovered this new member of the APOBEC family (Rogozin *et al.*, 2005). For the moment, the role of A4 is still unclear but its expression is primarily observed in mice testis, more particularly in round spermatids. A possible role is suggested in spermatogenesis (Hogg, 2010).

APOBEC3 proteins constitute the largest subfamily. It includes seven of the eleven members, which are: APOBEC3A (A3A), APOBEC3B (A3B), APOBEC3C (A3C), APOBEC3DE (A3DE), APOBEC3F (A3F), APOBEC3G (A3G) and APOBEC3H (A3H). All A3 genes are forming a cluster on chromosome 22 and are organized in tandem. The expression pattern varies among A3 members. The expression is tissue and cell-type specific. However, cells are not restricted to the expression of a single A3. Each cell and tissue express a range of proteins at different levels. Since some members of the A3 subfamily have 2 deaminase domains (A3B, A3DE, A3F, A3G) while others bear only one (A3A, A3C, A3H), they have different sizes. A3 proteins are unique to placental mammals but their numbers vary among phylogenetic clades (Salter *et al.*, 2016; Smith *et al.*, 2012). A unique A3 gene is found in rodents, cats, pigs and sheep. Two A3 genes are found in cows, three in dogs and horses, seven in primates and up to eighteen in bats (OhAinle *et al.*, 2006). Polymorphism is found in the human population and the different haplotypes can show different activity against their target (Salter *et al.*, 2016).

Table 1: Size, subcellular localization, number of deaminase domains, substrates of binding and deamination activity and favored motifs of APOBEC proteins. Size is expressed in number of amino acids. N is for nuclear; C is for cytoplasmic. The number in parentheses represents the number of catalytically active domains. The table compiles information from the text.

	Size	Subcellular localization	Deaminase domains	Binding activity	Deamination activity	Favored motif
AID	198	N/C	1	ssDNA/RNA	ssDNA/RNA	5'-WRC-3'
A1	215	N/C	1	ssDNA/RNA	ssDNA/RNA	5'-AC-3'
A2	224	N/C	1	?	?	?
A3A	199	N/C	1	ssDNA/RNA	ssDNA/RNA	5'-TC-3'
A3B	382	N	2 (1)	ssDNA/RNA	ssDNA	5'-TC-3'
A3C	190	N/C	1	ssDNA/RNA	ssDNA	5'-TC-3'
A3DE	366	C	2	ssDNA/RNA	ssDNA	5'-TC-3'
A3F	373	C	2 (1)	ssDNA/RNA	ssDNA	5'-TC-3'
A3G	384	C	2 (1)	ssDNA/RNA	ssDNA/RNA	5'-CC-3'
A3H	182	N/C	1	ssDNA/RNA	ssDNA	5'-TC-3'
A4	367	?	1 (0)	?	?	?

Functions of APOBEC3 subfamily

APOBEC3 members are primarily effectors of the immune system. Their role is to restrict exogenous viruses and endogenous retro-elements. The independent emergence of multiple subfamily members may have been selected in response to the diversity of these pathogens. Each A3 protein is more or less specialized against a given type of virus or retroelement (Ratcliff & Simmonds, 2021).

Retroelements, i.e., mobile genetic sequences that can copy and paste themselves elsewhere in the genome, could actually be the original targets of the A3 proteins. The retroelements are called 'retro' because they rely on reverse transcription to copy-paste. Among the retroelements, we differentiate the LTR-containing retroelements (endogenous retroviruses and retrotransposons) from the non-LTR-containing ones (the LINEs and SINEs). Importantly, the increase of A3 sub-members is associated with the reduction of retroelements activity in the human genome (Smith *et al.,* 2012). Thus, in mice, where a single A3 is present, retroelements are still active in their genome while they are not in humans. In addition, mice are 35 times more susceptible to diseases related to LINE-1 insertions (LINE-1 meaning Long Interspersed Nuclear Element). The mouse genome actually carries 50 to 60 times more active LINE-1 than the human genome (Esnault *et al.,* 2006; Schumann, 2007). In humans, 46% of the genome originates from inactivated transposable elements (Biémont & Vieira, 2006). The expansion of A3 proteins is thought to have provided protection against the genotoxicity of these elements. Due to their ability to self-copy, move and randomly insert themselves, they could be the source of many genomic instabilities (Salter *et al.,* 2016). Furthermore, all A3s have demonstrated the ability to restrict mobile elements. However, the APOBEC signature related to the transition from cytosine into uridine is not found in all types of transposable elements (Ratcliff & Simmonds, 2021). In fact, APOBEC proteins exhibit other mechanisms than deamination to restrict mobile genetic elements. Indeed, deamination is not the only tool of APOBEC proteins and restriction also occurs through deamination-independent events. Just as A3A can restrict LINE-1 through deamination-dependent mechanisms, A3G can sequester RNA from SINEs in cytosolic HMMs. Similarly, A3F can associate with the LINE-1 ORF2 protein to interfere with their retrotransposition as well as in that of SINEs since they are dependent on it (Chiu *et al.,* 2006; Richardson *et al.,* 2014; Stenglein & Harris, 2006).

The first function described for A3s were nonetheless viral restriction. In 2002 while studying HIV-1 mutants, an antiviral factor associated with HIV restriction was identified. That cellular actor was suspected to be the causal agent responsible for many guanosine-to-adenosine (G-to-A) transitions in the HIV provirus. Viruses affected by these hypermutation events were producing less infectious viral particles (Sheehy *et al.*, 2002). It turned out to be A3G (known as CEM15 at that time). Although APOBEC proteins deaminate cytosine into uridine, a strong bias of G-to-A mutations is found in the HIV-1 genome. Since HIV-1 is a retrovirus, one of the first steps of its life cycle is the reverse transcription of its positive-sense mRNA (+mRNA) into cDNA. A3G can deaminate the newly-synthesized complementary negative-strand ssDNA. According to the base complementary, the C-to-U mutation will be fixed as G-to-A given that this strand will serve as a template for the synthesis of the positive cDNA strand. These deamination events can affect more than 10% of the guanosine residues in the HIV-1 genomes (Harris & Dudley, 2015; Mangeat *et al.*, 2003; Zhang *et al.*, 2003). These are called hypermutations and can disrupt proper viral replication. The majority of hypermutation events actually occur during the following cellular infection. When an infected cell produces new virions, A3G associates with the HIV genome and the nucleocapsid domain (NC) of the Gag protein. During the budding of the mature HIV virion, A3G proteins are incorporated in the viral capsid. A3G will be able to directly act on the nascent cDNA during the subsequent infection. Thus, a large part of the proviruses that will be integrated will in fact be defective because of the abundance of mutations such as aborted start codons or nonsense mutations. Moreover, a fraction of the proviruses will be degraded even before their integration. Indeed, upon detection of uridine, DNA repair enzymes like the uracil-DNA glycosylase-2 (UNG2) will cleave the N-glycosidic bond and excise uracil from the sugar-phosphate DNA backbone. It leads to the introduction of abasic sites and the subsequential degradation of the strand by apurinic-apyrimidinic endonucleases (APE) (Chiu & Greene, 2009)

Deaminase-independent mechanisms are also involved in HIV-1 restriction. Indeed, A3G is able to bind RNA and to inhibit HIV-1 infectiousness. This capacity allows A3G to bind and to physically block the reverse transcription of the HIV genome (Chiu & Greene, 2009). It has also been observed that A3G is able to bind $tRNA^{Lys3}$, which is necessary for the priming of the reverse transcription, but also to disturb the removal of the tRNA primer during the positive-strand synthesis leading to the formation of abnormal 3' positive cDNA ends (F. Guo *et al.*, 2006; Mbisa *et al.*, 2007). A3G can also associate with HIV proteins required for its integration, such as the integrase or the reverse transcriptase, further altering its viral activity (Figure 9) (Chiu & Greene, 2009; Luo *et al.*, 2007).

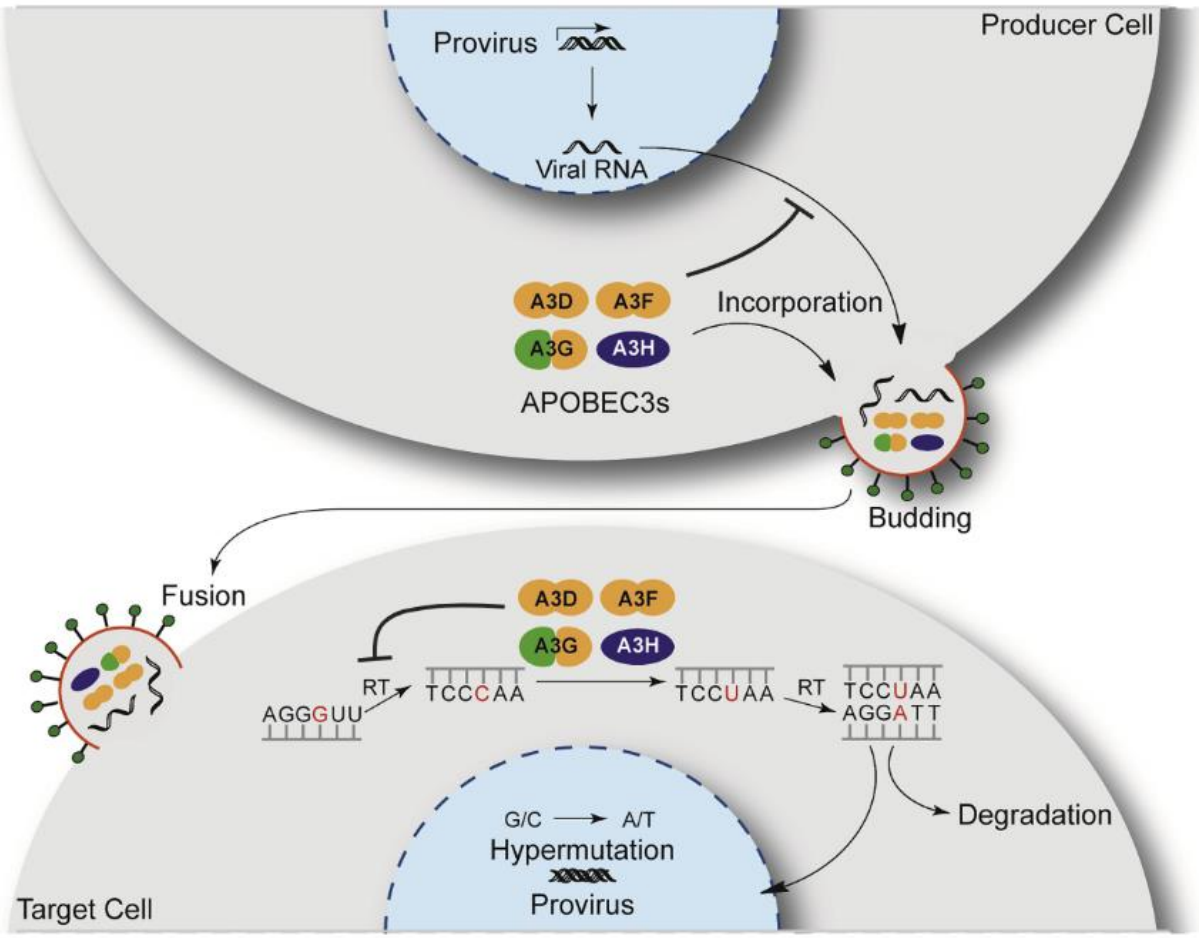

Figure 9: APOBEC proteins can restrict HIV through deaminase-dependent and independent mechanisms. APOBEC proteins hypermutate HIV genome leading to defective progeny or direct degradation. APOBEC proteins physically impede replication mechanisms (modified from Olson *et al.*, 2018).

Although A3G is the most effective actor against HIV-1, it turned out that the other members of the A3 subfamily can also display anti-HIV activity. Albeit to a lesser degree, A3F, A3DE, haplotype II, V and VII of A3H and one haplotype of A3C (I188) can also restrict HIV-1. (Delviks-Frankenberry *et al.*, 2020) They all have the ability to introduce C-to-U mutations and to bind RNA, making them ideal weapons against HIV-1. However, the main mechanism involved in HIV-1 restriction may vary among the A3s. For example, HIV-1 restriction by A3G primarily occurs through deaminase-dependent mechanisms, whereas A3F mostly uses deaminase-independent mechanisms (Mbisa *et al.*, 2010). As a result, the level of induced G-to-A mutations varies among A3s. Furthermore, since the preferential deamination motif is different between A3G (5'-CC-3') and the other A3 members (5'-TC-3'), it provides easier access to other viral sequences than those targeted by A3G. Altogether A3 proteins form a squadron of antiviral weapons to hamper viral replication (Delviks-Frankenberry *et al.*, 2020).

This ability of the A3 proteins to induce hypermutations and to impair the viral machinery is obviously not restricted to HIV-1. The A3 proteins notably target other reverse transcribing viruses. G-to-A hypermutations were actually observed in HIV-1 and the avian spleen necrosis virus (a gammaretrovirus) even before the involvement of A3s in HIV restriction was elucidated. In the end, other retroviruses from different families such as the Rous sarcoma virus (RSV) (alpha-retrovirus), the mouse mammary tumor virus (MMTV) (beta-retrovirus), the murine leukemia virus (MLV) (another gamma-retrovirus), the human T-lymphotropic virus (HTLV-1) (delta-retrovirus), foamy viruses (spumavirus) but also other lentiviruses such as the

equine infectious anemia virus (EIAV) or the simian immunodeficiency virus (SIV) were also found to be restricted by A3 proteins (Delviks-Frankenberry *et al.,* 2020; Smith *et al.,* 2012).

Although retroviruses have been extensively studied, A3s are actually able to restrict numerous groups of viruses. Indeed, double-stranded DNA (dsDNA) viruses such as the Herpes simplex virus 1 (HSV1), the Epstein-Barr virus (EBV) or the Human papillomavirus (HPV) and reverse transcribing viruses like the hepatitis B virus (HBV) have been shown to be hypermutated by A3 proteins. For example, both HBV negative and positive ssDNA strands are being deaminated and mutated by A3B, A3C, A3F et A3H. Although APOBEC proteins are known to target single-stranded nucleic acids, dsDNA viruses pass through a single-stranded phase during their life cycle. They then expose their genome for APOBEC proteins intervention, allowing dsDNA viral editing. Similar to retroviruses and retroelements, many of the antiviral mechanisms are driven by deaminase-independent events (Harris & Dudley, 2015; Ratcliff & Simmonds, 2021; Salter *et al.*, 2016). Such events are found in the enterovirus A71 (EV-A71) but also in parvoviruses such as in the adeno-associated virus (AAV) or the minute virus of mice (MVM). Indeed, deaminase-dead A3A is still able to restrict these two parvoviruses (AAV and MVM) despite the absence of deaminase activity (H. Chen *et al.,* 2006; Z. Li *et al.,* 2018; Narvaiza *et al.,* 2009)

The case of RNA viruses is particular because A3s, excluding A3A and A3G, are primarily known to only deaminate ssDNA substrates. Nonetheless, all A3s have shown the ability to bind RNA, opening up a field of possibilities related to deaminase-independent restriction mechanisms. A3G has been shown to restrict the measles virus (MV), the human respiratory syncytial virus (hRSV) and the mumps virus (MuV) (all of which are RNA viruses) but the underlying mechanisms and the question of the involvement of deamination-related effects are still unknown (Fehrholz *et al.,* 2012). On the other hand, the restriction of the Human coronavirus NL63 (HCoV-NL63) has been recently studied and it seems that A3C, A3F and A3H can restrict the HCoV-NL63 in a deamination-dependent manner. The authors could nonetheless not confirm the mechanism related to the viral restriction (Milewska *et al.,* 2018).

A3 proteins provide two ways to fight viruses: deamination-dependent mechanisms and their hypermutation events; and deamination-independent mechanisms that will clutter several actors of the viral infectious cycle. The two responses work in concert to slow down replication and alter the infectivity of virions.

Co-evolution and escape strategies

The relationship between hosts and pathogens is a perpetual arms race. Indeed, if the activity of A3G on HIV-1 could be discovered, it is because the study was conducted on HIV-1 mutants. This mutant was lacking an HIV-1 accessory protein, the virion infectivity factor (Vif). This viral protein is actually responsible for the degradation of the APOBEC proteins, therefore the ability to restrict the HIV-1 virus was obscured by the Vif protein (Sheehy *et al.,* 2002). In the course of evolution, each host and each pathogen evolve jointly to gain the upper hand. Thus, over time, viruses have co-evolved and developed many strategies to evade or even to counteract our defense mechanisms.

Vif is one of these actors able to counter the APOBEC response and prevent HIV to be restricted. First, Vif is able to prevent A3G from being encapsidated. A3G proteins actually need to oligomerize and to form homodimers to be packaged into mature virions. Vif binds and sequesters A3G proteins hindering the incorporation (Mariani *et al.,* 2003). But Vif is especially known to induce the degradation of A3G proteins by hijacking the cellular proteasomal

pathway. Vif triggers the A3G polyubiquitylation (but also its own) by recruiting the Cullin 5-complex, an E3 ubiquitin ligase. This complex includes the core-binding factor (CBFβ), the elongin B/C and the RING-box subunit 2 (RBx2). The mediated polyubiquitylation leads to A3G degradation by the host 26S proteasome (Figure 10) (Y. Guo *et al.,* 2014; Kobayashi *et al.,* 2005). Vif can also reduce the intracellular level of A3G by directly inhibiting the translation of A3G mRNA. Vif binds the 5' untranslated region (5' UTR) of A3G mRNA and prevents its translation (Guerrero *et al.,* 2016). In addition, it seems that Vif promotes the formation of HMMs, the inactive ribonucleoprotein complex composed of A3s and RNA. In the absence of Vif, HMMs are also present, but A3 proteins sequester viral RNAs within (Goila-Gaur *et al.,* 2008). A3G is actually not the only target of Vif: A3C, A3DE, A3F and A3H are also bound by Vif (Harris *et al.,* 2012). Altogether, these multiple layers of protection allow HIV-1 to remain infectious despite the presence of antiviral actors such as A3G.

Obviously, Vif is not the only example of anti-APOBEC proteins and many Vif-like proteins are found in other lentiviruses such as the simian immunodeficiency virus (SIV), the feline immunodeficiency virus (FIV), the bovine immunodeficiency virus (BIV), the maedi-visna virus (MVV) or the caprine arthritis encephalitis virus (CAEV) (Ratcliff & Simmonds, 2021). The relation between such Vif-like proteins and A3s is very specific. Consequently, the range of action of these Vif-like proteins is limited to the A3s of its host. For example, the human A3G is not sensitive to SIV Vif-like protein (Bogerd *et al.,* 2004; Mangeat *et al.,* 2004). In contrast, lentiviruses such as the MLV or the EIAV do not carry Vif equivalent (Smith *et al.,* 2012).

Foamy viruses also encode an accessory protein to prevent APOBEC antiviral actions, Bet. On the contrary to Vif, Bet does not trigger A3G proteasomal degradation but will rather hamper A3G encapsidation. Bet binds A3G which impairs its homodimerization and thus the packaging into mature virions. Indeed, a large part of A3s effects comes from their ability to be released with virions and to directly act in the subsequent infection. Bet could also sequester A3G into static complexes that will therefore not be incorporated during the virion budding (Jaguva Vasudevan *et al.,* 2013; Löchelt *et al.,* 2005). In the same way, the HTLV-1 resistance mechanism also involves a lower level of packaged A3G (Figure 10). However, in this case, no accessory protein is needed to prevent A3s actions. The HTLV-1 NC domain of the Gag protein contains a particular C-terminus extension. This motif disrupts the binding of A3G to the NC domain, which is required for A3G encapsidation. The exact process is not fully understood but the extension could be involved in the alteration of the structure or the modification of the RNA binding properties of the NC domain. Although different, both the underlying mechanisms result in less A3G being incorporated into virions (Derse *et al.,* 2007).

The EBV, a herpesvirus responsible for mononucleosis and certain lymphomas, is being targeted by A3B. To defend itself, EBV uses two mechanisms linked to an accessory protein: BORF2. This actor is able to bind and to stoichiometrically inhibit A3B deaminase activity. In addition, BORF2 relocalizes A3B into perinuclear and cytoplasmic bodies, keeping its viral genome in a safe place (Figure 10). The Kaposi's sarcoma herpesvirus (KSHV) also encodes a BORF2-like protein, ORF61 (Cheng *et al.,* 2019). In addition, KSHV tends to remain latent in B cells, however the expression of AID leads to the activation of its lytic cycle. To avoid this, KSHV encodes two miR that will target the 3' UTR of AID mRNA (Bekerman *et al.,* 2013).

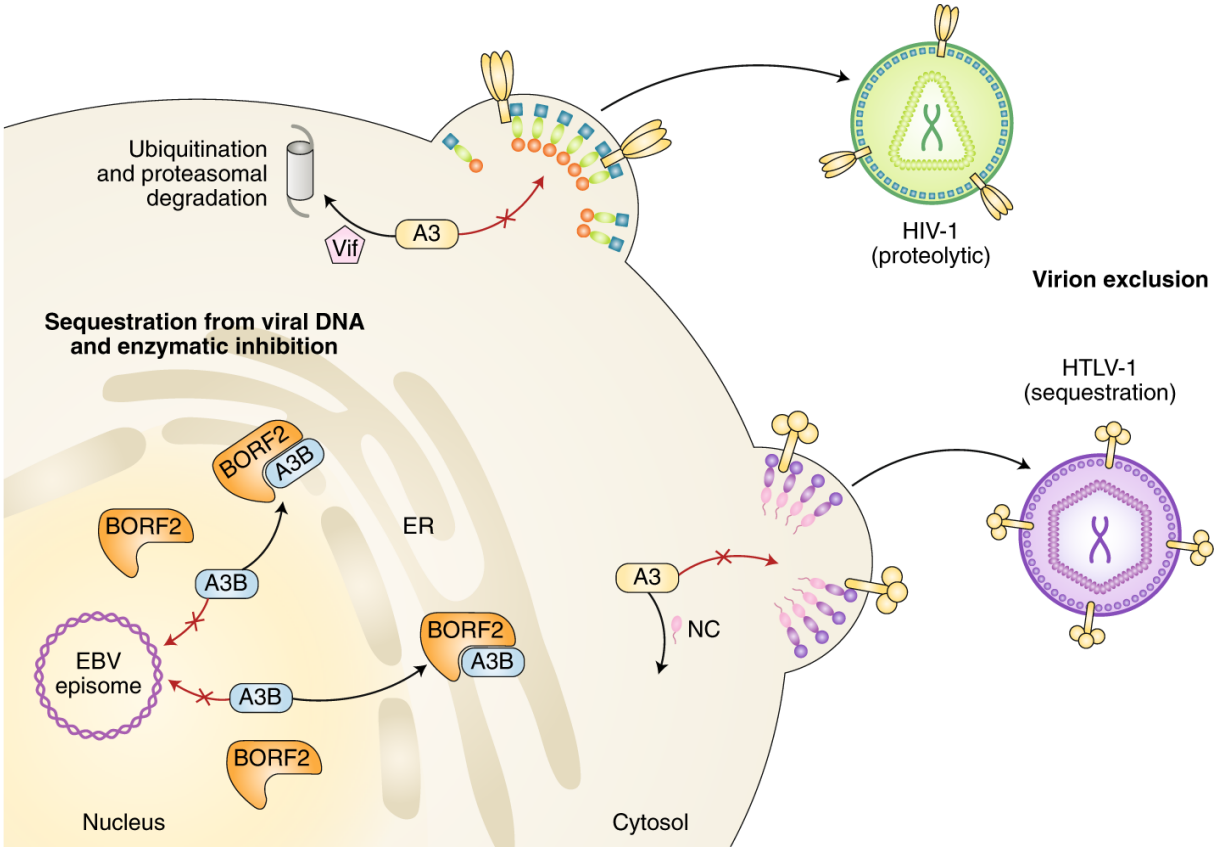

Figure 10: Virus developed escape mechanisms against APOBEC proteins: HIV Vif protein induces the proteasomal degradation, HTLV-1 excludes APOBEC proteins from the budding site, EBV BORF2 sequesters A3B away from its replication site (Malim & Pollpeter, 2018).

Besides these active mechanisms driven by accessory proteins, some viruses have evolved to reduce their exposure towards the A3s. Hence, some viruses show a reduced frequency of deaminase-targeted sites in their genome. Indeed, the selection pressure exerted by the deamination effect leads to the selection of genome depleted in motifs such as 5'-TC-3', the favored motif of most APOBEC proteins. It can be observed in BK polyomavirus or HPV (Verhalen *et al.*, 2016; Warren *et al.*, 2015).

Finally, some viruses have accommodated to APOBEC proteins and have domesticated their effects. This allows either to increase the selection of motif depleted-genome but also to increase their genetic diversity and thus promote their evolution (Kim *et al.*, 2014; R. Wang *et al.*, 2020). This effect could allow viruses to escape the immune system as well as selecting drug-resistant viral strains (Hernandez *et al.*, 2020; Neogi *et al.*, 2013). It has been observed that BK polyomavirus uses that scheme and upregulates A3B activity to shape its genome (Verhalen *et al.*, 2016). Indeed, when the virus succeeds in antagonizing the deaminase proteins, the virus can control the number of mutations to such a level that it becomes advantageous for them (Kim *et al.*, 2014; R. Wang *et al.*, 2020).

This diversity of approaches shows the fine balance between hosts and pathogens and the importance of APOBEC proteins in viral restriction.

SARS-CoV-2

Coronaviruses are a large group of widely spread viruses. Their name comes from the Latin word *corona*, which means crown. Indeed, when observed by electron microscopy they remind

us of a solar corona. This aspect comes from the presence of protruding spike peplomers (Haque *et al.,* 2020). Although certainly known since the 20's it is only in the 60's that they were described (Estola, 1970; Kahn & McIntosh, 2005). The first appearance of the name coronavirus dates from 1968 and the nomenclature is accepted in 1971 by the International Committee on Taxonomy of viruses (Lalchhandama, 2020; "Virology: Coronaviruses," 1968). Coronaviruses belong to the *Nidovirales* order which includes fourteen families, the major ones being *Coronaviridae, Arteriviridae* (Equine arteritis virus), *Mesoniviridae, Roniviridae* (Okavirus) and *Tobaniviridae* (Torovirus and Bafinivirus). *Coronaviridae* are split in *Orthocoronavirinae* and *Letovirinae*. *Orthocoronavirinae* are subdivided in four genera: *Alphacoronavirus, Betacoronavirus, Deltacoronavirus* and *Gammacoronavirus* (ICTV, 2020).

Coronaviruses can infect cells from the intestinal and the respiratory systems in diverse animals like birds, bats, mice, whales, minks or dogs. Some coronaviruses can cause severe disease in livestock and have significant economic impacts. (Hasöksüz *et al.,* 2020; Hu *et al.,* 2021). Alpha- and Betacoronavirus are restricted to mammalian hosts contrary to Delta- and Gammacoronavirus that infect a large range of hosts. Until December 2019, six coronaviruses were known to infect and be transmitted in the human population. The human coronavirus (HCoV) – NL63 and the HCoV-229 are alphacoronavirus, HCoV-HKU1 and HCoV-OC43 are betacoronavirus. These four coronaviruses are seasonal viruses that are known to circulate in the human population during the winter season. They cause cold-like symptoms and mild upper respiratory tract infections. Although, severe cases are occasionally seen in young children and the elderly. In 2002 and 2012, the 5th and the 6th human coronaviruses emerged. These two viruses, respectively the severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV), are highly pathogenic and infect the lower respiratory tract. People infected are susceptible to lung injuries and to develop severe and life-threatening respiratory complications (Hasöksüz *et al.,* 2020; V'kovski *et al.,* 2021). These two betacoronavirus are responsible for two major outbreaks and highlighted the threat of zoonotic diseases. Indeed, these two viruses would be of animal origin and would have succeeded in crossing the species barrier. Bats are known to be coronavirus reservoirs and intermediate hosts as the Himalayan palm civet (SARS-CoV) and the Arabian camel (MERS-CoV) could be the gateway that allowed the adaptation and the passage to humans. In late 2019, in China, the 7th human coronavirus, the SARS-CoV-2, emerges and causes the third human coronavirus outbreak. This new virus was so named because its genome is 85% similar to that of SARS-CoV. However, the SARS-CoV-2 virus far exceeds its betacoronavirus counterparts whether in terms of infected people, deaths, geographic spread or economic impact. (Baggen *et al.,* 2021; Haque *et al.,* 2020). This new virus is causing the global pandemic of coronavirus disease 2019 (COVID-19). This emerging disease has already infected more than 250.000.000 people and killed 5.000.000 of them (*WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data*, 2021).

The SARS-CoV-2 is a betacoronavirus. These viruses are enveloped non-segmented positive-sense single-stranded RNA viruses. The viral particles have a diameter of about 60 to 100 nm. The genome is ~30.000 bases long, 5' capped and 3' polyadenylated in addition to be flanked by UTR, at both 5' and 3' ends of 265 and 337 nucleotides, respectively. These UTR will serve as a *cis*-acting secondary structure for RNA synthesis (V'kovski *et al.,* 2021). The genome encodes at least eleven open reading frames (ORF). On the 3' side of the genome, the four major structural proteins are encoded: the nucleocapsid (N) protein, the membrane (M) protein, the envelope (E) protein and the Spike (S) protein (Figure 11). All these structural proteins are essential for the morphogenesis of the virions but they all have their particularities. The N protein coats the genome and packages the RNA to form the helical ribonucleocapsid. In addition, through its ability to bind RNA, the N proteins maintain the genome associated with

the replication/transcription complex (RTC). The N protein is also an IFN inhibitor and a repressor of the RNA interference pathway in other coronaviruses such as the SARS-CoV (Y. Chen *et al.*, 2020). The M protein is a component of the viral envelope and plays multiple roles in the virus assembly and budding. The M protein shapes the virions by interacting with the N protein and promotes its curvature. It will also turn the host membrane network into virions-producing factories. By binding multiple SARS-CoV-2 proteins, the M protein reassembles them at the budding site: the endoplasmic reticulum – Golgi intermediate compartment (ERGIC) lumen. Thanks to this, the different components of the future virions will be in the right place at the right time. It includes the S protein that is retained and accumulated in the ERGIC through an interaction between the M protein and its cytoplasmic C-terminus tail (Boson *et al.*, 2021; Y. Chen *et al.*, 2020; Neuman *et al.*, 2011). The E protein is the smallest structural protein. As the M protein, the E protein retains the S protein in the ERGIC by hijacking the cell secretory pathway. This process will also help in virions release (Boson *et al.*, 2021; Y. Chen *et al.*, 2020). Through its hydrophobic transmembrane domain, E proteins are able to self-assemble in the host membranes to form pentameric ion-channels known as "viroporins". These ions channels are associated with the pathogenesis of the SARS-CoV-2 (Nieto-Torres *et al.*, 2014). E proteins are also able to disrupt cellular tight junctions, increasing viral dissemination (De Maio *et al.*, 2020). The S protein is a glycoprotein and defines the virus tropism. By binding to its receptor, the angiotensin-converting enzyme 2 (ACE2), the protein allows entry into the cell. The protein forms homotrimers and each unit are made of two subunits. The subunit 1 is the surface unit and initiates the infection by interacting with ACE2. The subunit 2 is a transmembrane unit that promotes the fusion of the host and viral membranes. To mediate the entry into the host cell, the two subunits must be cleaved by a host protease, such as the transmembrane protease serine 2 (TMPRSS2) or the furin (Hoffmann *et al.*, 2018; Hoffmann, Kleine-Weber, & Pöhlmann, 2020; Hoffmann, Kleine-Weber, Schroeder, *et al.*, 2020).

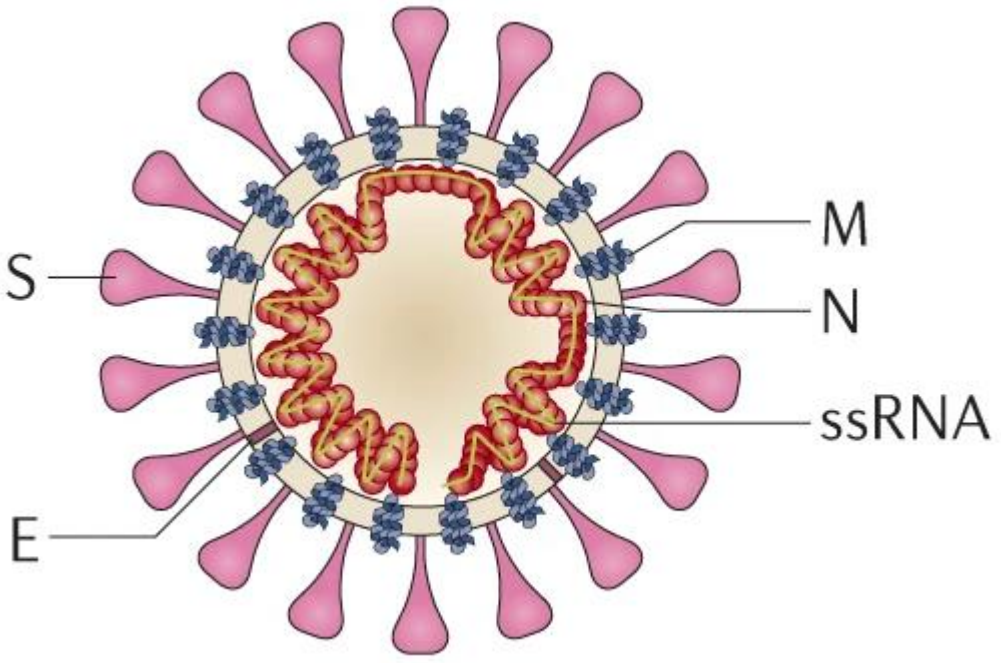

Figure 11: Schematic representation of SARS-CoV-2 virion. Major structural proteins and genome are represented. The viral particle includes spike (S), envelope (E), membrane (M), nucleocapsid (N) and the ssRNA genome (V'kovski *et al.*, 2021).

The 5' side of the genome is dedicated to non-structural proteins (nsp). This part of the genome encodes two overlapping ORF: ORF1a and ORF1b. When translated these two ORFs will give

rise to two polyproteins: pp1a and pp1ab. They are then co- and post-translationally processed into sixteen nsp (Figure 12a), fifteen of which are involved in the establishment of the replication/transcription complex (RTC), the protein complex that will support RNA synthesis. The nsp2-11 will promote a cellular environment conducive to viral replication while the nsp12-16 are directly involved in enzymatic reactions. Nsp1 is not directly involved in viral replication but rather blocks the host cell translation. It induces cellular mRNA degradation and leaves the field free for viral mechanisms. It also impedes the host cell immune response by interfering with the IFN signaling (Y. Chen *et al.*, 2020; Haque *et al.*, 2020). Some nsp functions remain unclear. Known functions of nsp are presented in Table 2. We can note the presence of an exoribonuclease equipped with a proofreading function. Indeed, due to the large size of its genome, such a function is essential to maintain genomic integrity (V'kovski *et al.*, 2021). In SARS-CoV mutants, the absence of this function considerably increases the mutation rate and impairs its growth (Eckerle *et al.*, 2010). Putative ORF are found on the 5' side of the genome and encodes non-essential accessory proteins (ORF3a/b/c/d, ORF6, ORF7a, ORF7b, ORF8 and ORF9b/c, ORF10). Their functions remain unclear but are certainly crucial for in vivo infection (Chan *et al.*, 2020).

Table 2: Putative functions and size of SARS-CoV-2 proteins sorted by their position in the SARS-CoV-2 positive-sense genome (Chan *et al*., 2020; Haque *et al*., 2020; *Severe Acute Respiratory Syndrome Coronavirus 2 | SWISS-MODEL Repository*, 2021).

Name	**Putative function**	**Amino acid size**
pp1a/b	replicases	7096
nsp1	host translation inhibition – suppression of immune response	180
nsp2	modulation of host cell survival signaling pathway	638
nsp3	protease – RTC and DMV formation	1945
nsp4	double membrane structures formation	500
nsp5	polyprotein processing	306
nsp6	double membrane structures formation – vesicular trafficking	290
nsp7	primase – complex with nsp8	83
nsp8	primase – complex with nsp7	198
nsp9	ssDNA/RNA binding	113
nsp10	mRNA cap methylation – cofactor for nsp14 and nsp16	139
nsp11	no function known	18
nsp12	RNA-dependent RNA polymerase	932
nsp13	RNA helicase	601
nsp14	proofreading exoribonuclease – N7-guanine methyltransferase	527
nsp 15	uridylate-specific endoribonuclease	346
nsp16	ribose 2'-O-methyltransferase	298
S	virus entry	1273
ORF3a	viroporin formation - lysosomal trafficking	275
ORF3b	interferon antagonist	22
ORF3c	no function known	41
ORF3d	no function known	57
E	structural protein	75
M	structural protein	222
ORF6	disruption of nuclear imports – interferon pathway inhibition	61
ORF7a	suppression of host tetherin and small interfering RNA	121
ORF7b	no function known	43
ORF8	host immune response modulation	121
N	structural protein	419
ORF9b	mitochondrial-associated innate immune response inhibition	97
ORF9c	no function known	73
ORF10	no function known	38

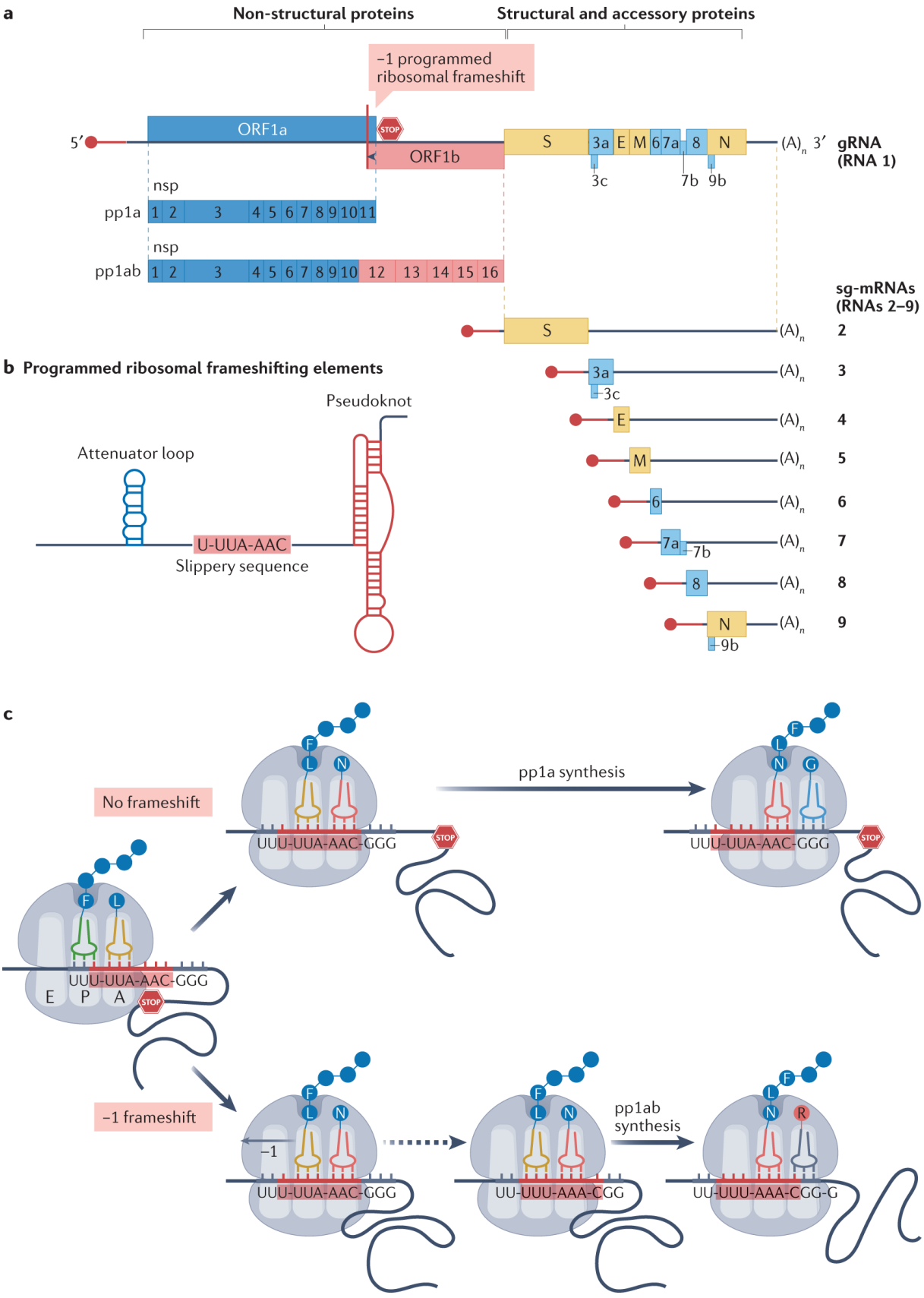

Figure 12: SARS-CoV-2 genome representation and regulation of gene expression. A. Genome organization of SARS-CoV-2, genome and subgenomic RNA are 5' capped and 3' polyadenylated. ORF giving birth to non-structural proteins are represented in blue and red (pp1a and pp1ab, respectively). Yellow represents ORF encoding structural proteins and light blue represents ORF encoding accessory proteins. B. The ribosomal frameshift is programmed on the slippery sequence thanks to interactions between ribosome and RNA structures. While translating the mRNA, the unfolding of the pseudoknot slows down

the ribosome and induces tension in the RNA template leading to a -1 frameshift. The attenuator loop modulates frameshifts frequency. C. Model for pp1a and pp1ab polyproteins translation. No frameshift leads to the reading of a stop codon and the synthesis of pp1a, frameshift allows the ribosome to continue the translation and leads to pp1ab synthesis (Malone *et al.*, 2021).

The SARS-CoV-2 is primarily a pathogen of the respiratory system and infects the epithelial cells of the upper respiratory tract. Once in the body, the virus binds ACE2 and starts its life cycle (Figure 13). The ability of a virus to infect a cell depends strongly on its ability to bind to its receptor. ACE2 is highly expressed in the lungs but not exclusively, the protein is also present in the intestines, kidneys, brain or heart. Therefore, SARS-CoV-2 can infect tissues other than the lungs. In addition, it is also possible that other receptors than ACE2 may promote the entry of the virus, increasing the range of host cells. When encountering its cellular host, the trimer of S proteins mediates the entry of the virus. The S protein binds ACE2 through its subunit 1 and the protein is cleaved by a host protease (such as TMPRSS2 or furin) and reveals the fusion peptide of the subunit 2. This fusion peptide allows the membranes to fuse. The virus can also enter the cell through endocytosis. In this case, the fusion is driven by the acidic pH and the endosomal cathepsins. Once in the cytoplasm, the N proteins release the +RNA genome, which can be translated as it is into pp1a and pp1ab polyproteins. Pp1a is directly translated from the ORF1a while pp1ab comes from a programmed -1 ribosomal frameshift (PRF) on the short overlapping sequence between ORF1a and ORF1b. This PRF is set up through a "slippery sequence" (5'-U UUA AAC GGG-3') (Figure 12b). During the translation of ORF1a, the ribosome can shift its reading frame by moving one nucleotide backward. This event is induced by the presence of a three-stemmed pseudoknot structure that interacts with the ribosome and triggers the shift (Figure 12c). ORF1a is translated 1,5 to 2 times more than ORF1b. This rate is regulated by the nascent protein and RNA structures such as an attenuator loop and the unfolding of the pseudoknot (Figure 12b). The deregulation of this ratio leads to defective viruses. From the two translated polyproteins, 16 nsp are released by the proteolytic activity of two protease domains, the papain-like protease within nsp3 and the chymotrypsin-like protease within nsp5. Nsp1-11 originate from pp1a and pp1ab produce nsp1-10 and 12-16. Once released, nsp1 takes control of the translation machinery and nsp 2-11 prepare the environment by transforming the membranes into virus factories. Indeed nsp3,4,6 induce the formation of organelles dedicated to the viral replication. From the endoplasmic reticulum (ER) membranes, perinuclear double-membrane structures will be formed: the double-membrane vesicles (DMV), the convoluted membranes (CM) and the double-membranes open spherules (DMS). These structures serve as anchors for the RTC but also provide a great niche for viral replication. Thus housed, the virus can gather all the molecules and actors needed for its replication. It will also avoid the innate immune receptors found in the cytosol. The nsp 12-16 are directly involved in enzymatic reactions of the RTC. Viral genome synthesis begins with the replication of a full-length negative-sense strand RNA by the RdRp (nsp12), aided by two primases (nsp7-8). This negative-sense genomic RNA will serve as a template for new positive-sense strands, which will either enter a new translation-transcription cycle or be incorporated into new virions. The production of structural and accessory proteins depends on the synthesis of subgenomic RNA (sgRNA). This process consists of a discontinuous transcription, which is characteristic of *Nidovirales.* In the 3' part of the genome, the region that encodes the structural proteins, a particular motif is found: 5′-ACGAAC-3′. This motif is placed in front of most ORF and is called the transcription regulatory sequence (TRS), more particularly TRS body (TRS-B). Another TRS is found (TRS-L) next to a leader sequence 70 nucleotides away from the 5' end. Nine TRS are found in the SARS-CoV-2 genome (eight TRS-B and one TRS-L). During the transcription, when the RdRp encounters TRS-B, it stops and resumes in front of the TRS-L. It leads to the production of a series of negative-sense sgRNA (-sgRNA) that will be transcribed into +sgRNA (Figure 14). Although theoretically polycistronic, they concretely encode one cistron. It is believed that only the first encountered ORF is translated. In addition to the eight

canonical sgRNA (S, 3a, E, M, 6, 7a, 8, N), non-canonical transcripts have been identified (3b/c/d, 7b, 9b/c, 10) and may derive from RNA recombination. Newly synthesized RNA leaves the replication organelles to be translated. Once translated, structural proteins join the ERGIC by transiting through the ER and meet positive-strand genomic RNA to be assembled. The budding takes place in the ERGIC lumen and the new virions will leave the cell through exocytosis (Baggen *et al.*, 2021; Haque *et al.*, 2020; Malone *et al.*, 2021; V'kovski *et al.*, 2021).

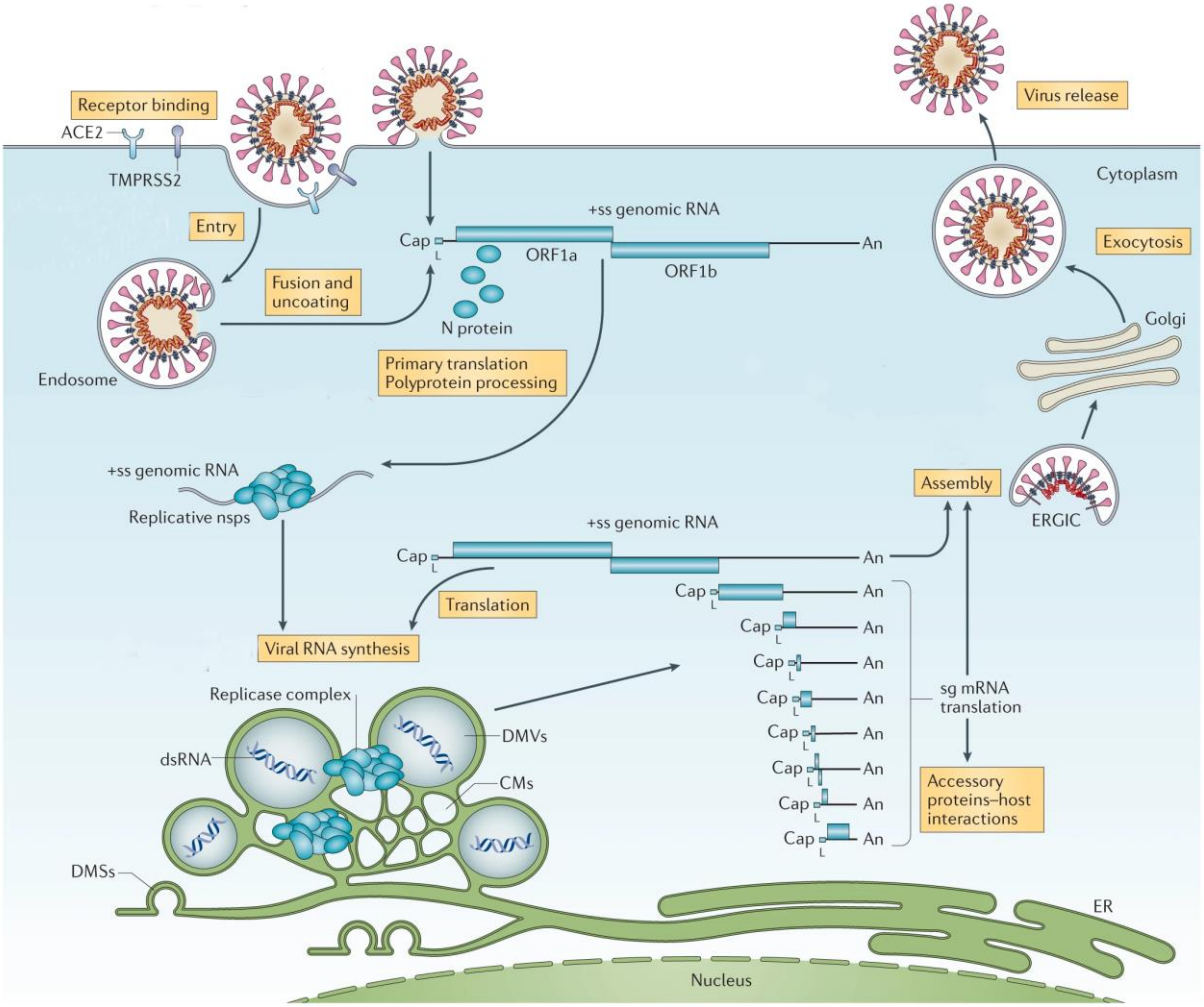

Figure 13: SARS-CoV-2 life cycle. The virion enters the cell either by endocytosis or by interaction with ACE2 through its spike protein. Cellular proteases at the membrane surface cleave the bond and allow the virions to enter the cell. Once in the cytoplasm, the viral particle is dismantled and the polyproteins are processed and produce sixteen non-structural proteins that will form the replication/transcription complex (RTC). Viral replication takes place in double-membrane structures (DMV, DMS, CM) set up by the virus and hijacked from the endoplasmic reticulum. This niche provides a protective micro-environment for the viral replication. Transcription of the viral genome leads to the production of genomic and subgenomic mRNA. Genomic RNA is either translated into non-structural proteins or packaged into new virions. Subgenomic mRNA are translated into structural and accessory proteins. Structural proteins transit to the ERGIC via the ER and are encapsidated with genomic RNA to be released (V'kovski *et al.*, 2021).

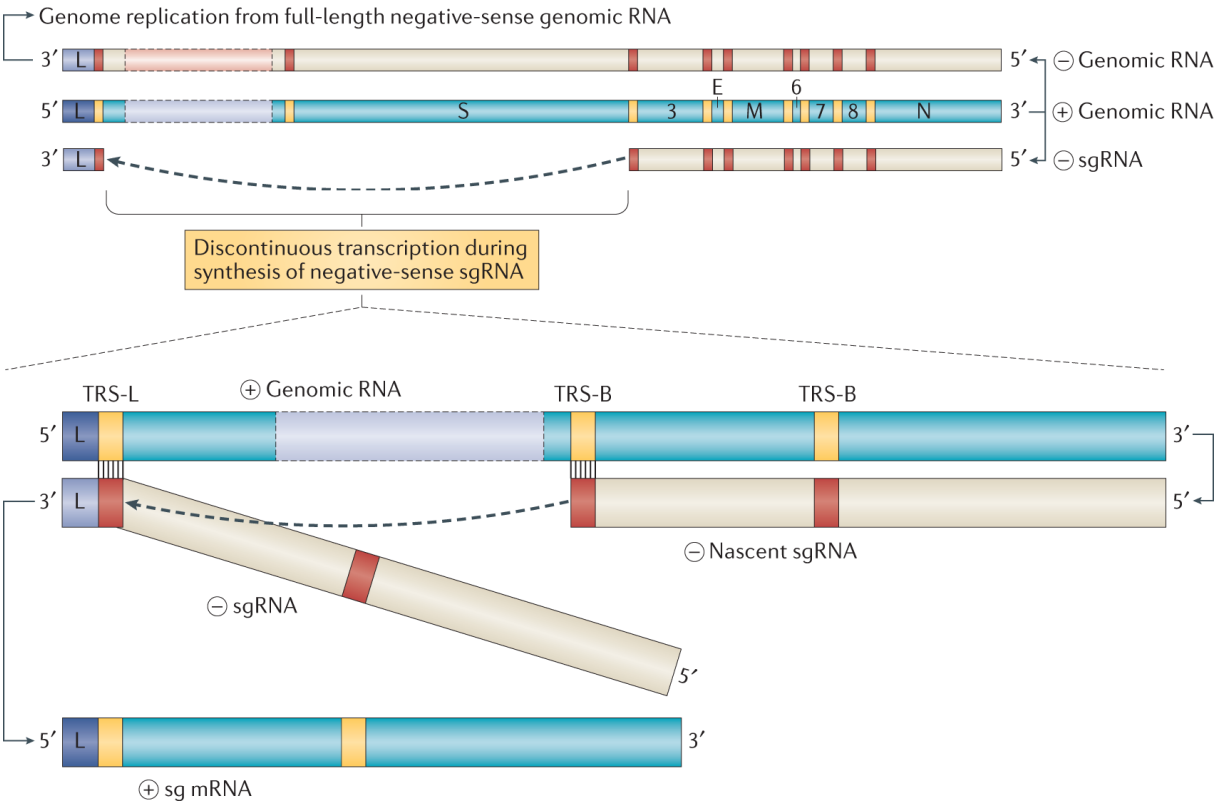

Figure 14: Coronavirus continuous and discontinuous RNA synthesis. Positive-sense gRNA gives rise to full-length negative-sense gRNA and negative-sense sgRNA (-sgRNA) by a discontinuous transcription process. These -sgRNA will serve as templates for positive-sense sgRNA. After copying TRS-B from the -sgRNA the RTC can relocalize on the template to jump on TRS-L found near the 5' end. This event gives rise to sgRNA that will be used for the production of structural and accessory proteins (V'kovski *et al.*, 2021).

APOBEC mutational signature

When an event changes the distribution of nucleotides, a bias is created. If this bias is not corrected, it becomes the imprint of this event. When the same event is associated with a particular mutagenesis process, it is called a mutational signature (Kockler & Gordenin, 2021). These signatures can have multiple origins such as unfaithful DNA replication or repair mechanisms, genotoxins such as ultraviolet radiation or alkylating agents but also enzymatic editing processes (Alexandrov *et al.*, 2013; Forbes *et al.*, 2017). For example, the adenosine deaminase acting on RNA (ADAR) deaminates adenosine into inosine on double-stranded RNA. Because inosine is read as guanine during the RNA replication by the RdRp, the genome is imprinted with A-to-G substitution (Samuel, 2011). Another example is the reactive oxygen species (ROS) that oxidize guanine into 7,8-dihydro-8-oxo-20-deoxyguanine (oxoguanine) and leads to G-to-U substitution. Oxoguanine pairs with adenine and ultimately leads to uridine mutations (David *et al.*, 2007). The APOBEC proteins are one of these enzymatic editing processes that can bias the nucleotide distribution.

APOBEC proteins exert a strong selection pressure on viruses. Consequently, viruses are forced to continuously evolve to escape this response. It leads to the selection of active escape mechanisms, but also leads to the reduction of targeted site. Indeed, the APOBEC threat led to the selection of non-lethal mutations that can avoid the targeting by APOBEC proteins. Throughout evolution, APOBEC proteins have driven virus evolution and shaped their genome (Poulain *et al.*, 2020). When analyzing virus genome, an underrepresentation of some motifs is

found. They can have multiple origins. In many viruses, the 5'-TC'-3' is less frequently found than expected. This motif is the favored target of APOBEC deaminases. This is, among others, the case for EBV, KSHV and BK polyomavirus (Martinez *et al.*, 2019; Warren *et al.*, 2015). Retroelements also show an underrepresentation of APOBEC motifs. Less 5'-TC-3' and 5'-CC-3' motifs are found on their negative strand, respectively the favored motifs of A3s and A3G (Anwar *et al.*, 2013). The signature does not always appear as such in certain viruses. For example, in HIV, the APOBEC signature takes the form of a G-to-A substitution. Indeed, the DNA intermediate of HIV is negative-sensed. Therefore, during the replication, the C-to-U substitution will be fixed as G-to-A in the positive sense genomic RNA (Jern *et al.*, 2009). This underrepresentation is called the APOBEC footprint. Obviously, the A3 footprint is only found in viruses that infect hosts encoding A3 genes. Avian and fish viruses do not bear the A3 signatures whereas it is present in primate viruses (Poulain *et al.*, 2020).

Recently Poulain *et al.*, scanned the genome of 870 human viral species for A3 footprints. The result showed that 175 (22%) species are footprinted. It includes many double-stranded DNA viruses such as the *papillomaviridae* (*alpha-*, *beta-* and *gammapapillomaviridae*) and the *polyomaviridae* (*beta-* and *deltapolyomaviridae).* This also includes single-stranded DNA viruses such as the autonomous *parvoviridae* or single-stranded RNA viruses such as the *coronaviridae*, of which B19 erythroparvovirus and HCoV-HKU1 are particularly imprinted. Besides these globally impacted viruses, the A3 footprint can be found in particular regions of the viral genome. For example, in *herpesviridae* and *adenoviridae*, A3 footprint is only observed in sequences that will serve for the initiation of the viral replication. It is assumed that this particularity is due to the way the virus replicates (Poulain *et al.*, 2020).

Footprinted viruses are nonetheless still sensitive to A3 restriction as most of the deamination events on the remaining motifs lead to mutations that are non-synonymous (Poulain *et al.*, 2020).

Evidence for SARS-CoV-2 restriction by APOBEC proteins

In less than two years, SARS-CoV-2 has infected more than 250.000.000 people. Because of this, a huge amount of data has been collected. Thanks to this and to the next-generation sequencing, SARS-CoV-2 genomes have been extensively dissected throughout the COVID-19 outbreak (Ratcliff & Simmonds, 2021). One of the methods to study and to analyze the SARS-CoV-2 evolution is to compare viral genomes obtained from infected patients with a reference sequence. This reference sequence (MN908947) has been obtained the 5 January 2020 from one of the first infected patient in Wuhan (Wu *et al.*, 2020). Thanks to its early origin, it may be the closest to the original ancestor. It is composed of 32,08% U, 29,94% A, 18,37% C and 19,61% G (Matyášek & Kovařík, 2020). Thus, when comparing single-nucleotide variants (SNV) from latter infected people to the reference, the direction of the viral evolution can be observed. In this way, Mourier *et al.* observed three major distribution biases. He found ~37% of C-to-U transitions, ~18% of G-to-U transversions and ~11% of A-to-G transitions that they attributed to APOBEC proteins, ROS and ADAR proteins, respectively (Mourier *et al.*, 2021). In the same manner, Sadykov *et al.* analyzed ~22.000 SARS-CoV-2 sequences and observed comparable results, ~29% of C-to-U transitions, ~15% of G-to-U transversions and ~14% of A-to-G transitions (Sadykov *et al.*, 2021). More than 33.000 SARS-CoV-2 isolates were analyzed by Wang *et al*. and they measured the proportion of the 12 possible single-nucleotide polymorphisms (SNP). They observed ~24% of C-to-U transitions, ~13% of G-to-U transversions and ~15% of A-to-G transitions (R. Wang *et al.*, 2020). From these three studies, it clearly appears that the nucleotide distribution is biased between SARS-

CoV-2 isolates and the supposed original ancestor, particularly in C-to-U transitions. Several actors can shape viral genomes. Due to the nature of this transition, APOBEC proteins can easily be suspected as one of the actors responsible for such bias. In addition, these C-to-U transitions are mostly found in sites that contain the APOBEC favorite motifs (Di Giorgio *et al.*, 2020; Simmonds, 2020). Intriguingly, no APOBEC footprint is found in bat viruses. Yet they are the mammal whose A3 subfamily is the most diversified, with no less than 18 distinct A3 genes in Pteropid bats (Hayward *et al.*, 2018).

A3 footprint was already observed in the endemic human coronaviruses such as HCoV-OC43, HCoV-229E, HCoV-HKU1 and the HCoV-NL63 but not in zoonotic coronaviruses such as the MERS-CoV and the SARS-CoV (Poulain *et al.*, 2020). The absence of A3 footprint on such viruses could be explained by the fact that, unlike endemic coronaviruses, these zoonotic viruses have co-existed for a short time with the human population. Indeed, SARS-CoV and MERS-CoV have respectively infected 8.000 and 2.500 people in short epidemic events (Institut Pasteur, 2011; World Health Organisation, 2019). Although significant on the human scale, these events are relatively minimal on the evolutionary scale. This is another story for SARS-CoV-2 as the outbreak is still ongoing and the number of infected people, but also the geographical extent, has taken a whole new dimension. Consequently, as the epidemic progresses and human transmissions multiply, the cohabitation between our immune system and SARS-CoV-2 could lead to the emergence of the A3 footprint on the viral genome. Canonically, A3 proteins only deaminate DNA targets. However, it was recently observed that A3A and A3G could use RNA as substrate for deamination (Sharma *et al.*, 2019). This opens the door to the possible restriction of coronaviruses by APOBEC proteins. Furthermore, Milewska *et al*. showed that HCoV-NL63 could actually be restricted by APOBEC proteins in vitro. The associated mechanism has not been elucidated, it may rely on sole or both deaminase-dependent and -independent mechanisms (Milewska *et al.*, 2018). In fact, A3 signature have been found on the rubella RNA virus and the N protein of HCoV-229E, HCoV-NL63 and SARS-CoV have been shown to interact with A3 proteins (Perelygina *et al.*, 2019; S. M. Wang & Wang, 2009). Although no hypermutation events have been detected, deaminase-dependent restriction is still an option. Indeed, the genome of HCoV-NL63 virus is particularly low in cytosine (14%). This is possible that this virus has reached its maximum level of cytosine depletion. In this case, hypermutation events would be hidden by the selection of unaffected and therefore unmutated viruses. It is also possible that hypermutation events are actually repaired by the proofreading function of their exonuclease. In addition, when infecting APOBEC over-expressing cells, C-to-U point mutations have been observed (Milewska *et al.*, 2018).

For these multiple reasons: i. human endemic coronaviruses evolved under A3 selection pressure (Poulain *et al.*, 2020), ii. the mutations observed in SARS-CoV-2 have been attributed to the APOBEC enzymatic activity on the basis of the substitution types and the nucleotides context surrounding the mutated sites (Di Giorgio *et al.*, 2020; Mourier *et al.*, 2021; Sadykov *et al.*, 2021; Simmonds, 2020; R. Wang *et al.*, 2020), iii. in vitro experiments demonstrated that A3C, A3F and A3H can restrict the replication of the endemic alphacoronavirus HCoV-NL63 (Milewska *et al.*, 2018), iv. recent report showed APOBEC deamination on ribocytidine substrates (Sharma *et al.*, 2015, 2019). Based on these observations, we investigated the relationship between APOBEC proteins and SARS-CoV-2. In particular, whether SARS-CoV-2 replication induces the expression of A3 proteins and whether these proteins can restrict viral replication.

Materials and methods

Cell lines – Vero E6 (epithelial kidney cells from Cercopithecus aethiops) and the Vero E6 transduced to over-express human APOBEC3 (or GFP control) were propagated in Eagle's Minimum Essential Medium with EBSS and L-glutamine (EMEM; BioWhittaker) supplemented with 10% fetal bovine serum (FBS). Transduced Vero E6 cells are resistant to blasticidin.

HEK 293T (Human embryonic kidney 293T cells) cells were grown in Dulbecco's Modified Eagle Medium with 4,5 g/L D-glucose, L-glutamine and pyruvate (DMEM; Gibco) supplemented with 10% FBS.

HBEC3-KT (human bronchial epithelial cells) are immortalized with the human telomerase reverse transcriptase (hTERT) and the mouse cyclin-dependent kinase (CDK4). HBEC3-KT and transduced HBEC3-KT-ACE2 were grown in gelatinized flasks with keratinocyte-SFM (1X) (Serum-free medium; Gibco) with L-Glutamine complemented with keratinocyte-SFM Supplement: Human Recombinant Epidermal Growth Factor (EGF 1-53) & Bovine Pituitary Extract (BPE) (Gibco). HBEC3-KT are resistant to puromycin and geneticin. HBEC3-KT-ACE2 are resistant to puromycin, geneticin and hygromycin. HBEC3-KT-ACE2-ShA3G/ShSCr/A3G/GFP are resistant to puromycin, geneticin, hygromycin and blasticidin.

All cell lineages were incubated at 37°C with 5% CO2 in humidified incubator.

SARS-COV-2 production – Vero E6 cells were cultivated in EMEM supplemented with 2% FBS. When Vero E6 reached 80% confluence, cells were infected with SARS-CoV-2 with a multiplicity of infection of 0,01 and incubated for 4 days. Cell medium was centrifugated for 5 min at 400g and the supernatant was harvested and frozen at -80°C. Viral titer were quantified with a typical 50% Tissue Culture Infectious Dose (TCID50) assay. The SARS-CoV-2 strain used was isolated from a Belgian COVID-19 patient in March 2020 by Prof Piet Maes (KU Leuven). This isolate belongs to the original strain, clade G.

Lentiviral plasmid construction – The A3 lentiviral expression plasmids were based on a modified version of the pLenti4 backbone. The A3 CDS is cloned in an antisense orientation relative to the lentiviral genome and split by a sense-oriented intron as described by Law *et al.,* 2016. This construction allows conditional expression of the APOEBC3 protein in the transduced cells but not in the packaging cells. Indeed, A3s being antiretroviral proteins, they would corrupt the lentiviral particles if expressed in the HEK 293T cells. Cloning of the different fragments were made using the Gibson assembly technique (NEBuilder® HiFi DNA Assembly, New England Biolabs). The PCR primers used to amply the A3 CDS and the Beta-globin intron are reported in SI (Table S1 and Table S2). The constructions bear a blasticidin resistance gene. The respective APOBEC transgenes are under cytomegalovirus promoter control and are HA-tagged.

ShA3G (pSicoR_EF1α_mCherry_T2A_Blasti_shA3G) lentiviral plasmid were cloned using a modified version of pSicoR lentiviral vector. This vector encodes a puromycin resistance gene with a ribosome skipping sequence located after the mCherry reporter (pSicor-MS2). Two 5' phosphorylated page-purified oligonucleotides were annealed to form the shRNA A3G fragment. The two oligos (100µM) were incubated with CutSmart® buffer (10X) (New England Biolabs) at 95°C for 4 min, at 70°C for 10 min and slowly cooled down to 4°C (-1°C/1min). pSicoR backbone (pSicoR_EF1α_mCherry_T2A_Puro, already in use in the lab)

was restricted with Hpa1 (New England Biolabs) and Xho1 (New England Biolabs). Annealed oligos were restricted with the same enzymes and cloned in a previously opened backbone using T4 DNA ligase (10X) (New England Biolab). Puromycin resistance gene was replaced with a blasticidin resistance gene. Blasticidin resistance gene were amplified from pSicoR_EF1α_mCherry_T2A_Blasti_sh_scrambled (ShSCr) and restricted with Sma1 (Takara) and EcoR1 (Takara). The resulting plasmids were restricted with the same enzymes and blasticidin resistance gene fragment cloned within using T4 DNA ligase (New England Biolabs). Oligos used for shA3G transgene construction are reported in SI (Table S3).

ACE2 lentiviral plasmid were provided by Neville Sanjana (purchased on Addgene, plasmid #161758, pLenti-hACE2-hygro). The construction bears a hygromycin gene resistance. The ACE2 transgene is under eEF1α promoter.

The maps of the lentiviral plasmid are reported in SI (Fig S1-16). Bacterial transformations were made using One Shot™ TOP10 Chemically Competent E. coli (ThermoFisher Scientific). Plasmid purifications were made using NucleoSpin Plasmid, Mini kit for plasmid DNA (Macherey-Nagel). All enzymes and kits were used according to the manufacturer's instructions. Plasmids were verified by Sanger sequencing.

Lentivirus production – Lentiviruses for Vero E6 and HBEC3-KT transductions were produced by transfecting HEK 293T cells with the plasmid DNA constructs previously described. HEK 293T cells were transfected with lipofectamine 2000 (ThermoFischer Scientific) at 80% confluence with the "APOBEC" lentiviral plasmids (respectively A1, A2, A3B, A3DE, A3F, A3G, A3H and GFP constructs) or ACE2 lentiviral plasmid or "shRNA" lentiviral plasmids (shA3G, shSCr, A3G, GFP) and the helper plasmids psPAX2 (gifted from Didier Trono, Addgene plasmid # 12260), and pCMV-VSV-G (gifted from Robert Weinberg, Addgene plasmid # 8454). The medium was renewed 6 hours post-transfection (PI). Cells were maintained for 24 hours in DMEM supplemented with 10% FBS and lentiviruses were harvested from the medium and filtered with a 0,22 µm filter. The medium was renewed and lentiviruses were harvested a second time and filtered 24 hours after the first harvest.

Vero E6 and HBEC3-KT transduction – Vero E6 cells were transduced with previously produced "APOBEC" lentiviruses at 70% confluence. Filtered lentiviruses were added to each flask containing Vero E6 cells and 0.1% polybrene (Sigma-Aldrich) was added to the medium. Transduced cells were maintained in EMEM and selected with blasticidin (15µg/ml, InvivoGen) for 7 days to eliminate non-transduced cells. Intron removal was verified by PCR and sequencing.

HBEC3-KT were first transduced with ACE2 lentivirus in the same conditions but selected with hygromycin (3mg/ml). Proper transduction of ACE2 transgene was verified using fluorescence-activated cell sorting (FACS).

HBEC3-KT-ACE2 were transduced a second time with "shRNA" lentiviral vectors in the same conditions as Vero E6 cells.

Fluorescence-activated cell sorting – After antibiotic selection, about 10^6 cells (~1/16 of T175 flask) were harvested by trypsinization, centrifugated for 5min at 400g and the supernatant discarded. Cells were washed with PBS 10% FCS for trypsin neutralization, centrifugated for 5min at 400g and the supernatant discarded. Cells were resuspended in eBioscience™ IC Fixation Buffer (Invitrogen) incubated for 30min, centrifugated and the supernatant discarded.

Pellet is resuspended eBioscience™ Permeabilization Buffer (10X) (Invitrogen, 1/10 with deionized water), incubated for 30min, centrifugated and the supernatant discarded. Cells are resuspended in 100 µl PBS 10% FCS with anti-ACE2 antibody (Abcam, ab15348, 1%). The solution is incubated 90min and mixed each 30min with pipet up-and-down. 100µl of PBS 10% FCS is added, the mix is centrifugated and the supernatant discarded. Pellet is washed with 200 µl PBS 10% FCS, centrifugated and the supernatant discarded. Cells are resuspended in 100µl PBS 10% FCS with anti-rabbit Alexa 405 secondary antibody (Invitrogen, A-31556, 0,1%), and incubated 60min with pipet up-and-down each 30min. 100µl PBS 10% FCS is added, the mix is centrifugated and the supernatant discarded. Pellet is resuspended in 500µl FACSFlow™ Sheath Fluid (Becton, Dickinson and Company) and transferred in FACS tube for cytometry analysis.

SARS-CoV2 infection – One day before the Vero E6 infection, 20 000 cells were seeded in 96-well plates in 200ul of EMEM supplemented with 10% FBS. Cells were inoculated at a multiplicity of infection of 0,01 or mock-infected in EMEM supplemented with 2% FBS. Two hours later the medium was renewed with EMEM supplemented with 2% FBS. At 2-, 24-, 48-, 72-, 96-hours post-infection 100 µl of supernatant were harvested and frozen at -80°C for RNA extraction. Cell viability was assessed by luciferase assay as described thereafter. Triplicates were made for each condition

HBEC3-KT-ACE2 were seeded one day before infection in complemented keratinocyte-SFM (1X) medium to reach 60% cell confluence the infection day. Cells were infected with SARS-CoV2 at a multiplicity of infection of 0,5 or 5. The medium was renewed 3 hours later. At 6-, 48-, 96-hours post-infection 100 µl of supernatant were harvested and frozen at -80°C for RNA extraction. At 2- and 4- days post-infection (dpi), cells were harvested for protein and RNA extraction.

Cells were harvested by trypsinization, washed with Phosphate Buffer Saline without calcium or magnesium (BioWhittaker) (PBS) and resuspended either in 100 µl RIPA (50 mM Tris HCl, pH 8.0; 150mM NaCl; 1% NP-40; 0.5% sodium deoxycholate; 0.1% SDS) supplemented with proteases inhibitors cocktail 2% Complete, 10% Triton X-100 and 1% PMSF 1mM (cell fraction intended for Western blot) or in 100 µl HED buffer (20mM HEPES, pH7.4; 5mM EDTA; 1 mM DTT; 10% glycerol) supplemented with proteases inhibitors cocktail 2% Complete (cell fraction intended for deamination activity assay) or in 900 TRIzol™ Reagent (ThermoFisher Scientific) (cell fraction intended for RNA quantification).

Luminescent cell viability assay – Cell viability was determined using the CellTiter-Glo® Luminescent Cell Viability Assay (Promega) according to the manufacturer's instructions.

DNA extraction – After transduction, cells were verified by sanger DNA sequencing. Cells were harvested by trypsinization and resuspended in PBS. DNA extracted using NucleoSpin® Tissue Kit (Macherey-Nagel) according to manufacturer's instructions.

Protein extraction – Cells were transferred in sonication tubes and incubated on ice for 30 min. Cells were sonicated 5 cycles high-frequency 30sec ON / 30sec OFF (cell fraction intended for Western blot) or 15 cycles low-frequency 30sec ON / 30sec OFF (cell fraction intended for deamination activity assay) at 4°C with a Bioruptor® device (Diagenode). Products were centrifugated at 4°C for 15min at 14 000 rpm. Supernatants were harvested and quantified using Pierce™ BCA Protein Assay Kit (ThermoFischer Scientific) according to the manufacturer's

instructions. Quantifications were made using the Thermo Scientific Multiskan EX (ThermoFischer Scientific) microplate photometer reading at 550nm.

Western Blot – Before loading, samples were boiled for 5 min at 95°C. Depending on the protein size, 30 µg of proteins were loaded in 10% or 12% SDS-PAGE gel and migrated for 3 hours at 100V. Proteins in the gel were transferred to a PolyVinyliDene Fluoride membrane previously activated with methanol for 1 min. Proteins were transferring at 200V, 200mA, 50W at 4°C for 1h45. Membranes were blocked with TBS-T BSA 5% for 3 hours (Tris-buffered Saline, 0,1% Tween-20, 5% Albumin fraction V (Sigma-Aldrich)). Primary antibodies against HA-tag (Medical & Biological Laboratories Co., LTD., PM032, 0,04%), Hsp90 (Invitrogen, 2-2.2.14, 0,1%), A3C (Proteintech, 10591-1-AP, 0,2%), A3F (National Institute of Health, 5206-235-07, 0,005%), A3G (National Institute of Health, 5210-87-13, 0,1%) and A3H (National Institute of Health, P3A3-A10, 0,1%) were incubated overnight at 4°C on the membrane. Primary antibodies were diluted in TBS-T BSA 5%. The next day, membranes were washed 3 times for 10min with TBS-T. Secondary antibodies were added and incubated for an hour at RT. Polyclonal Swine Anti-Rabbit Immunoglobulins/HRP (Dako) was used as secondary antibodies against Hsp90-, A3C-, A3F- and A3G-focusing primary antibodies and Polyclonal Goat Anti-Mouse Immunoglobulins/HRP (Dako) were used against HA-tag and A3H-focusing primary antibodies. Both secondary antibodies were diluted in TBS-T BSA 5% at 0.05%. Membranes were washed 3 times for 15min with TBS-T. Membranes were revealed using the ImageQuant LAS4000 mini biomolecular imager (Ge Healthcare Life Sciences) with a cooled CCD camera. Pierce™ ECL Western Blotting Substrate (ThermoFischer Scientific) were used as revelation reagents for the 12% membrane and Hsp90. SuperSignal™ West Femto Maximum Sensitivity Substrate (ThermoFischer Scientific) were used as revelation reagents used for the 10% membrane. All membranes were incubated with their respective reagent for 3min.

Deamination activity assay – Protein extract was incubated overnight at 37°C with 1mM $ZnCl_2$, 1µl of Uracil-DNA glycosylase (New England Biolabs, 1%), 2 µl of UDG reaction buffer (New England Biolabs, 10X), 1µl of RNAse A (New England Biolabs) and the TCC-Cy5 DNA probe (5'-ATT ATT ATT ATT CCC AAA TGG ATT TAT TTA TTT ATT TAT TTA TTT-Cy5-3'). In the control condition, the protein extract is replaced by 1pmol of the pre-deaminated TUA-Cy5 DNA probe (5'-ATT ATT ATT ATT UAA ATG GAT TTA TTT ATT TAT TTA TTT ATT T-Cy5-3'). After incubation, 2µl of 1M NaOH are added and the mix is heated at 95°C for 10min. Products are mixed with formamide loading buffer and loaded in 15% urea-acrylamide gel. Probes migrated at 200V for 90min and were revealed using ImageQuant LAS4000 mini biomolecular imager (Ge Healthcare Life Sciences) with a cooled CCD camera.

RNA extraction – After harvesting, Vero E6 cells RNA was extracted and purified using TANBead® Nucleic Acid Extraction Kit according to manufacturer's instruction and frozen at -80°C.

After harvesting, HBEC-3KT-ACE2 cells-TRIzol mix was vortexed for 30sec and incubated for 5min. 200µl of chloroform was added, the mix was vortexed for 30sec, incubated for 5min and centrifugated for 10min at 12000g at 4°C. The upper phase of the trilayer mix was harvested. 500 µl of isopropanol is added to the harvested phase, the mix is vortexed for 30sec, incubated for 5min and centrifugated for 10min at 12000g at 4°C. The supernatant is discarded

and the pellet resuspended in 1ml ethanol 70%. The mix is centrifugated for 10min at 12000g at 4°C. Ethanol is discarded and pellet resuspended in deionized water. Purified RNA was frozen at -80°C.

Viral RNA quantification - RNA quantification was performed using Takyon™ One-Step No Rox Probe 5X MasterMix dTTP (Eurogentec) according to the manufacturer's instruction. Thermal cycler protocol and primers are reported in SI (Table S4).

Cellular RNA quantification – Purified cellular RNA was first retro-transcribed into cDNA using iScript™ cDNA Synthesis Kit (Bio-Rad). 500ng of RNA were mixed with 5x iScript Reaction Mix, iScript Reverse transcriptase and incubated for 5min at 25°C, 30min at 46°C and 1min at 95°C. Reagent volumes were adjusted according to the manufacturer'sinstructionsn. cDNA was quantified using Takyon™ No ROX SYBR 2X MasterMix blue dTTP (UF-NSMT-B0701, Eurogentec). Each sample was diluted 10, 30 and 90 times. APOBEC mRNA quantification was performed using appropriate reverse and forward primers (3µM of each), 3µl of cDNA and 5µl of 2X MasterMix for a total of 10 µl in each well. Thermal cycler protocol and primers are reported in SI (Table S5).

Results

Establishment of APOBEC-overexpressing Vero E6 cell lineages

The first step of our approach was to establish cell lines that overexpress the human APOBEC proteins and a GFP control cell line. For that, we transduced wild-type Vero E6 cellsµ (African Green Monkey kidney cell line highly permissive for SARS-CoV-2) with lentiviral vectors bearing the human APOBEC transgenes. Since APOBEC proteins have antiretroviral activity, a particular method had to be used. Indeed, the expression of APOBEC proteins in the lentivirus-producing HEK 293T cells could induce deleterious mutations for our lentivirus production. Therefore, we have used a conditional system where the APOBEC proteins will be only expressed in the transduced cells and not in the cells producing the lentiviral particles (i.e., the HEK-293T cells). As described by Law *et al.*, the transgene is inactive because the open reading frame is interrupted by a sense β-globin intron while the transgene is antisense oriented (Figure 15). The intron will be removed by splicing in the virus-producing cells and expressed only once reverse-transcribed and integrated into the transduced Vero E6 cells. Blasticidin has been used to ensure the selection of the effectively transduced cells. A cell line transduced with a GFP expression construct has also been made and will be used as control in our subsequent experiments.

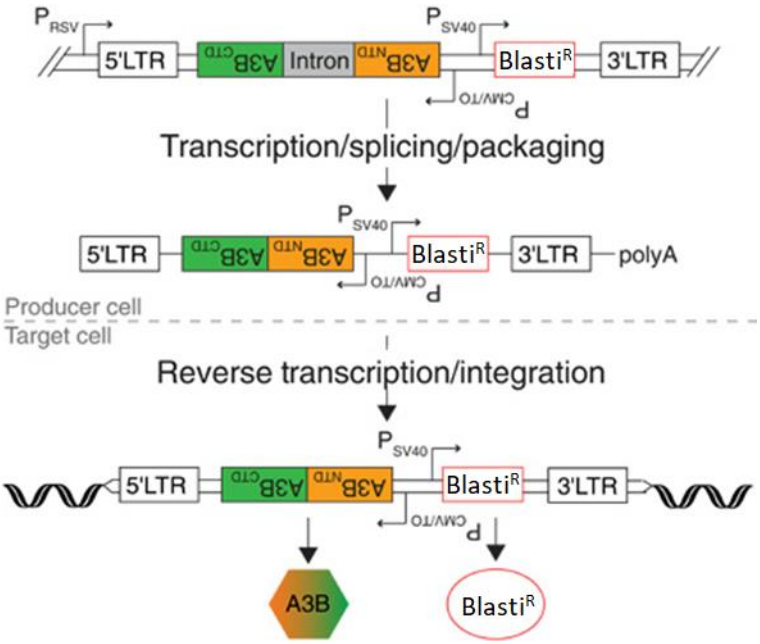

Figure 15: lentiviral construction approach for conditional APOBEC expression. P: promoter, RSV: Rous sarcoma virus, LTR: long terminal repeat, CTD: C-terminal domain, NTD: N-terminal domain, CMV: cytomegalovirus, SV40: simian virus 40, Blasti^R: Blasticidin resistance gene (modified from Law *et al.*, 2016*).*

APOBEC proteins are expressed in the established cell lineages.

We verified the expression levels of our HA-tagged APOBEC proteins by Western Blot using an anti-HA antibody. The expected sizes of the APOBEC proteins were the following: A1-HA is expected at 31kDa, A2-HA at 28kDa, A3A-HA at 26kDa, A3B-HA at 41kDa, A3C-HA at 25kDa, A3DE-HA at 49kDa, A3F-HA at 48kDa, A3G-HA at 49kDa and A3H-HA at 25kDa. Figure 16 showed that the A2, A3A, A3C and A3H Vero E6 cell lines do express a protein that is recognized by the anti-HA antibody and that has migrated at the expected size. We observed that the relative expression of the APOBEC proteins differs between the cell lines, A2 and A3H being more expressed than the A3A and A3C. We also observed that the A2 protein displayed different sizes. The A1 protein is not produced. The GFP protein is not HA-tagged and therefore no signal has been observed in the corresponding blot. We observed by fluorescent microscopy that the cells transduced with the GFP-expressing lentiviral vectors were efficiently transduced. Figure 17 showed that the A3B and A3C Vero E6 cells express a protein at the expected size. No clear expression has been observed for the A3DE and A3G Vero E6 cells. A band at the expected size has been observed for the A3F veroE6 cells albeit together with an unexpected shorter protein. Normalization is provided by the Hsp90 loading charge. To summarize, we validated the construction of the veroE6 cells overexpressing the A2, A3A, A3B, A3C and A3H. The A1, A3DE and A3G Vero E6 cells did not produce the expected protein whereas being transduced with a validated lentiviral vector (i.e., verified by Sanger sequencing). To understand the reasons for these failures, we chose to analyze the sequences of the integrated transgenes.

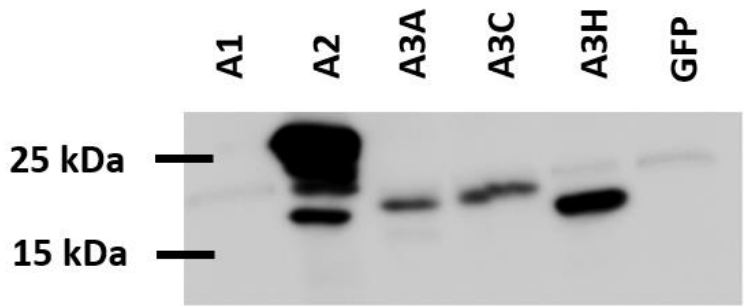

Figure 16: All APOBEC proteins are not expressed in the transduced Vero E6 cells. Assessment of APOBEC proteins in 12% SDS gel western blot. Hsp90 protein is used as a loading control.

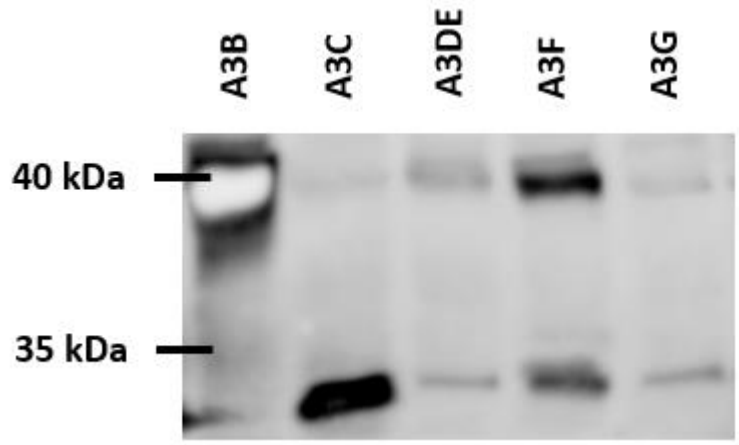

Figure 17: All APOBEC proteins are not expressed in the transduced Vero E6 cells. Assessment of APOBEC proteins in 10% SDS gel western blot. Hsp90 protein is used as a loading control.

Retaining of the β-globin intron is associated with the absence of expression of the APOBEC protein

We extracted and purified DNA from each lineage and conducted PCR spanning the transgene. If the intron is retained, the size of the amplified fragment is increased by 573bp (base pair) and will be identified by electrophoresis. A1 is expected at 910bp, A2 at 874bp, A3B at 1348 bp, A3C at 772bp, A3DE at 1260bp, A3F at 1321bp, A3G at 1354bp and A3H at 751bp.

A1 and A3DE are not correctly spliced and the intron is being retained. Two bands can be seen for A3F and A3G lineages (Figure 18). It seems that there is a heterogeneous cell population with some cells that were transduced with a lentivirus that has been correctly spliced and some with a lentivirus that retained the intron. A1, A3DE and A3G have been excluded as the splicing is not correct and no proteins were detected in the western blot (Figure 16-17-18). The PCR products were Sanger-sequenced to verify their integrity.

To summarize, we chose to assess the replication kinetics of the SARS-CoV-2 in the A2, A3A, A3B, A3C, A3F and A3H-expressing Vero E6 cells. New lentiviral constructions have been reconstructed for A3DE, A3F and A3G with a modified position of the intron within the coding sequence of the deaminase.

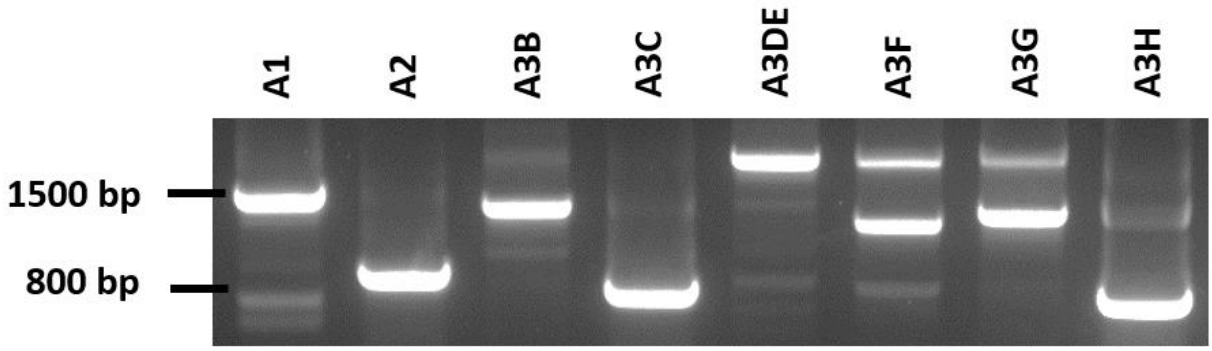

Figure 18: PCR screening of intron splicing. Primers were placed up and downstream of the transgene to monitor correct splicing. Unexpected bands are seen in A1, A3DE, A3F and A3H.

Exogenous expression of APOBEC in Vero E6 did not appear to modify the SARS-CoV-2 induced cell death

SARS-CoV-2 replication in Vero E6 cells induces cell death of the infected cells. Thus, measurement of cell viability allows an indirect estimation of viral replication. We therefore conducted a Luminescent Cell Viability Assay where the ATP is used as an indicator of cell viability. Luciferin is converted into oxyluciferin thanks to the luciferase provided. The reaction produces light which is measured by a luminometer. The luminescent signal is proportional to the amount of ATP as the reaction needs it and thus it reflects the number of viable cells.

We did not observe any obvious impact of the APOBEC expression on the cell viability. The viral replication induced a similar decrease of cell viability both in APOBEC-expressing, GFP and WT cell lines (Figure 19).

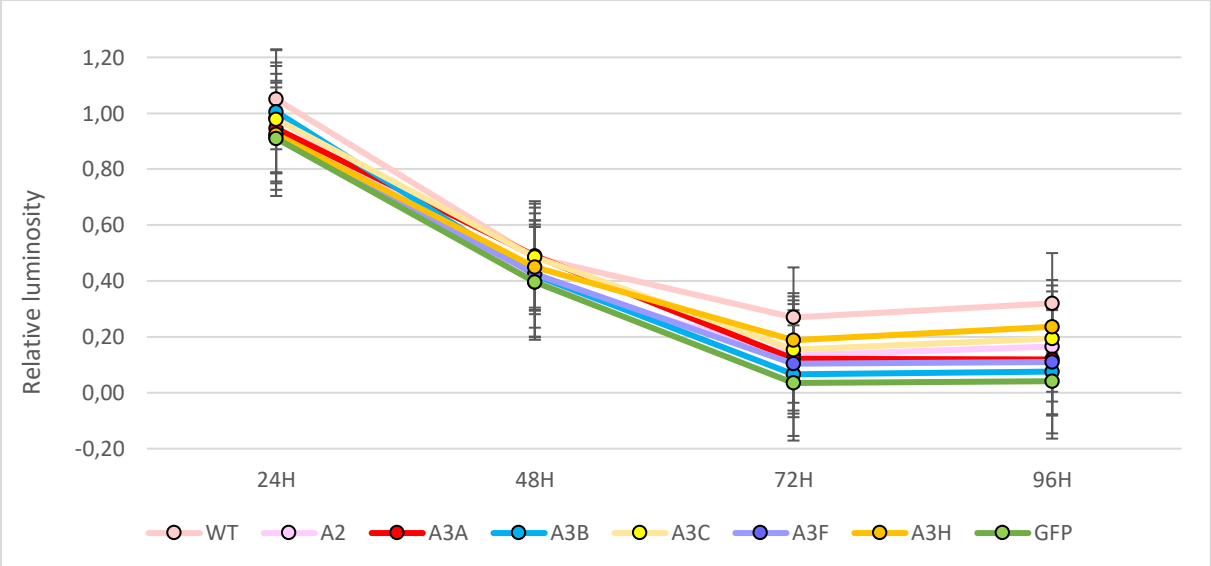

Figure 19: Cell viability assessment at 24-, 48-, 72- and 96 hours post-infection. Relative luminosity emitted by the 6 APOBEC-expressing, GFP and WT cell lineages after treatment is used to estimate cell viability.

Cell viability was also estimated by microscopy looking at cytopathic effects. We did not observe obvious differences in cell viability between the APOBEC-expressing cells lines compared to GFP and WT controls.

Virus replication is not impaired by APOBEC-overexpressing Vero E6

APOBEC overexpressing Vero E6 cell lines have been infected with SARS-CoV-2 in order to compare virus fitness. Kinetics of replication was assessed by quantification of the viral RNA by RTqPCR. To do this, we collected the virus genome from the supernatant of the infected cells at 24-, 48-, 72 hours post-infection. We then amplified a fragment of the E gene of the virus and reported the Cycle threshold (Ct) in figure 20. We did not observe any obvious impact of the APOBEC expression on the production of viral genomes in the culture supernatant.

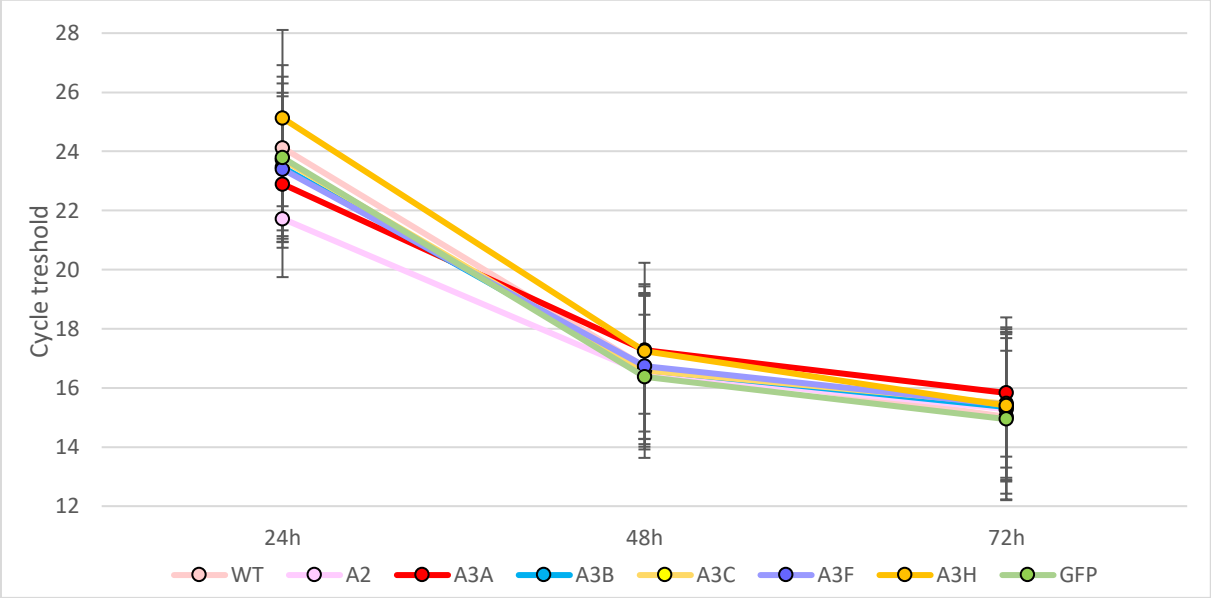

Figure 20: Measurement of Ct at 24-, 48-, 72- and 96 hours post-infection reflecting the kinetics of replication of SARS-CoV-2 infecting APOBEC- overexpressing cell lineages. SARS-CoV-2's E gene is used as target for probe hybridization.

Transduced HBEC3-KT cells express ACE2 protein

We changed our approach and chose to monitor the expression profile of APOBEC proteins upon SARS-CoV-2 infection of human bronchial epithelial cells. For that we used HBEC3-KT cells. These cells are primary cells and do not express ACE2 at this stage of differentiation. They are therefore non-permissive for SARS-CoV-2 infection. To overcome this, we genetically modified HBEC3-KT cells to allow the entry of the virus through the ACE2 protein. We transduced the HBEC3-KT cells with a lentiviral vector carrying the ACE2 cDNA to get the HBEC3-KT-ACE2 cells. We verified the ACE2 expression level by FACS. We observed that our HBEC3-KT-ACE2 homogeneously express high level of ACE2. Two cell lines are used for ACE2 level comparison: Vero E6 cells that constitutively express ACE2 and HBEC3-KT-ACE2 that have been transduced. Light blue (Vero E6) and orange (HBEC3-KT-ACE2) lines are control conditions, cells have only been incubated with the secondary antibody. It allows to monitor the unspecific binding of the secondary antibody. The blue line represents Vero E6 cells incubated with primary and then secondary antibodies. The Red line represents HBEC3-KT-ACE2 and have also been incubated with both antibodies. HBEC3-KT-cells express a higher level of ACE2 than the constitutive expression of Vero E6 cells.

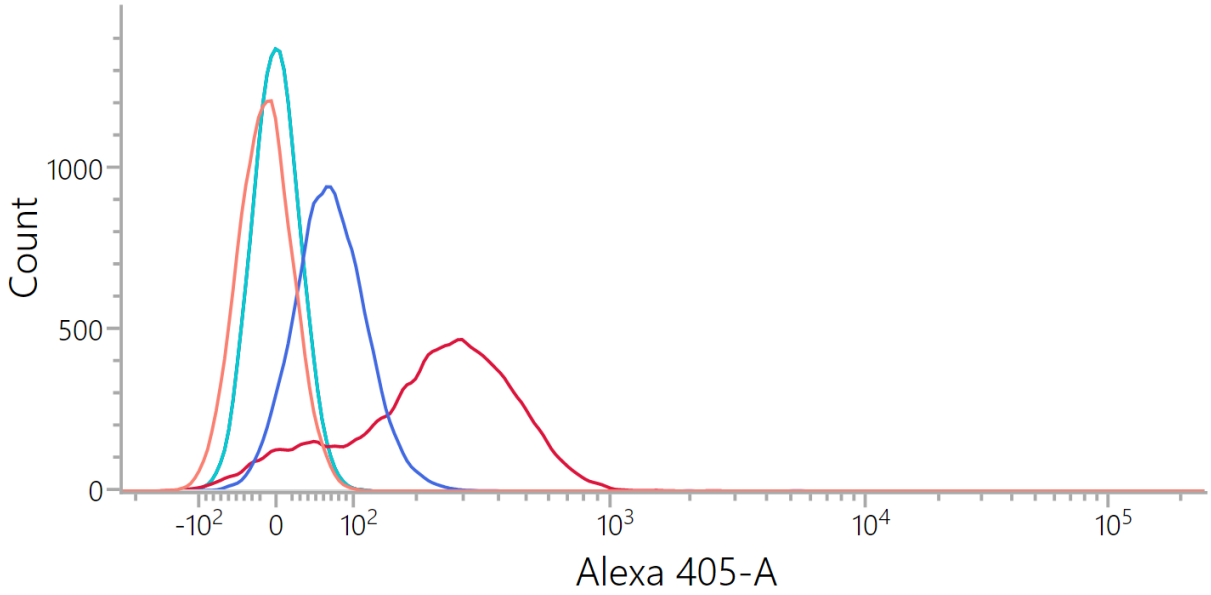

Figure 21: HBEC3-KT-ACE2 cells express a high level of ACE2 protein. Light blue and orange lines represent controls conditions where cells have only been incubated with secondary antibodies (Vero E6 and HBEC3-KT-ACE2, respectively). The blue line represents ACE2 level of Vero E6 and the orange line represents ACE2 level of HBEC3-KT-ACE2.

SARS-CoV-2 infection of HBEC3-KT-ACE2 cells leads to strong cytopathic effects

In order to assess the expression profile of APOBEC proteins upon SARS-CoV-2 infection, we grew HBEC3-KT-ACE2 to 60% confluence and infected them with SARS-CoV-2 at an MOI of 0,5 and 5. We observed cytopathic effects in a dose and time-dependent manner, indicative of the viral replication. This observation also allows us to confirm that the ACE2 transduction allows the entry of the virus in these previously non-permissive cells. Non-infected (mock) cells show no cytopathy and continue to grow, eventually forming a complete cell mat. Cells infected at a multiplicity of infection of 0,5 show moderate cytopathic effects at 2 days post-infection. The effect is increased at 4 days post-infection. Cells infected at a multiplicity of infection of 5 show strong cytopathic effects at 2- and 4-days post-infection. Half of the cells have detached and are floating in the supernatant. The difference observed between mock and infected cells shows the effects of SARS-CoV-2 infection both in a dose and time-dependent manner (Figure 22).

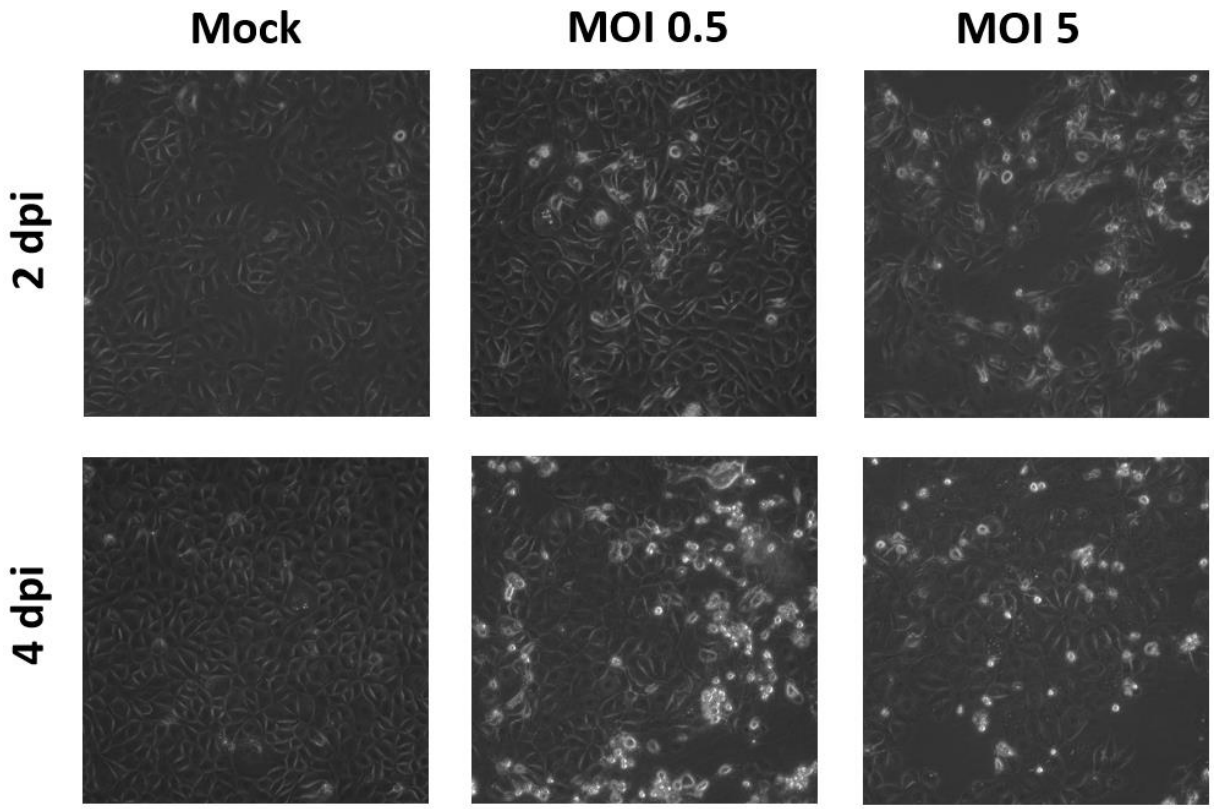

Figure 22: Infected HBEC cells show mild to strong cytopathic effects in a dose and time-dependent manner. HBEC cells were grown in an atmosphere containing 5% CO_2 at 37° and infected with SARS-CoV-2 at an MOI of 0,5 and 5. Cytopathic effects were observed with an inverted microscope 2- and 4-days post-infection.

Viral load is increased upon SARS-CoV-2 infection of HBEC3-KT-ACE2 cells

In addition to the observation of cytopathic effects, we measured the viral production in the supernatant. For this purpose, we harvested supernatant from the infected cell culture medium at 6-, 48- and 96-hours post-infection. We extracted RNA from this supernatant and evaluated viral growth by quantitative RT-qPCR. Viral replication was assessed by amplification of the E gene of SARS-CoV-2 and is reflected by the quantity of viral genomes. Ct are reported in figure 23. In both infection doses, viral load increases over time and shows that the virus actually replicates.

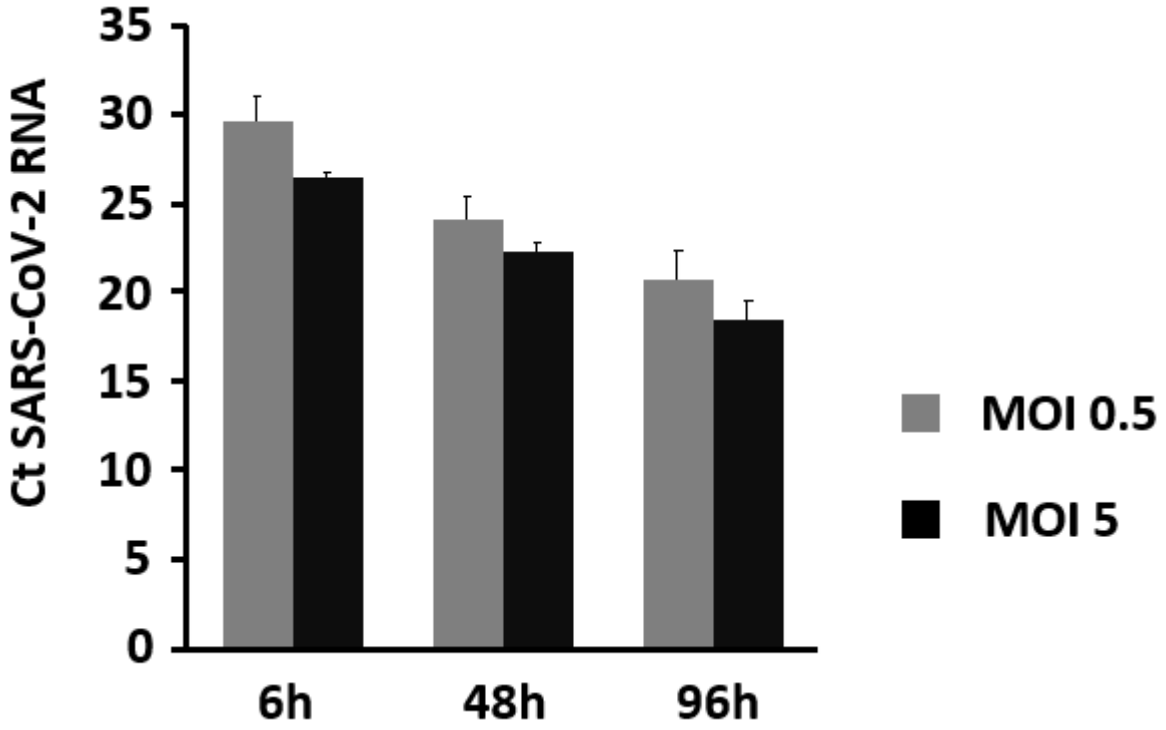

Figure 23: Viral load increases over time in the infected cell culture supernatant. Viral replication was measured at 6-, 48- and 72-hours post-infection and reported as Ct values for the SARS-CoV-2 E gene.

Infection of HBEC3-KT-ACE2 with SARS-CoV-2 alters the expression profile of APOBEC3 mRNA

We investigated the expression profile of the A3s mRNA following SARS-CoV-2 infection of our HBEC3-KT-ACE2 bronchial epithelial cells. Once grown and infected, we harvested the infected cells and extracted cellular RNA. We retro-transcribed RNA into cDNA and ran quantitative PCR in order to evaluate the expression level of the A3s mRNA. We also assessed the abundance of the mRNA of three housekeeping genes: hypoxanthine-guanine phosphoribosyltransferase (HPRT), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and TATA-binding protein (TBP). These genes are not involved in mechanisms linked to viral infection and should therefore not vary upon viral infection. For this reason, these genes will be used for the normalization of the APOBEC induction assessment. We chose to compare the level of expression of A3s mRNA between non-infected cells and infected cells. Infection was done in triplicate and the mean fold change is plotted in figure 24. RTqPCR analysis reveals that APOBEC expression level greatly differs among them. A3B, A3F and A3H are significantly more expressed in cells, respectively, 4 days post-infection at an MOI of 5, 4 days post-infection at an MOI of 0,5 and 2 days post-infection at an MOI of 0,5. A3G is significantly more expressed 2 days post-infection both at an MOI of 0,5 and 5, significantly more expressed 4 days post-infection at an MOI of 0,5 and highly significantly more expressed 4 days post-infection at an MOI of 5. Other conditions show no significant fold change between non-infected cells and infected cells (Figure 24). A3A and A3DE expressions were not detectable. Individual measurements of A3s mRNA expression level relative to housekeeping genes in each replicate are plotted and available in supplemental data (Figure S17-19).

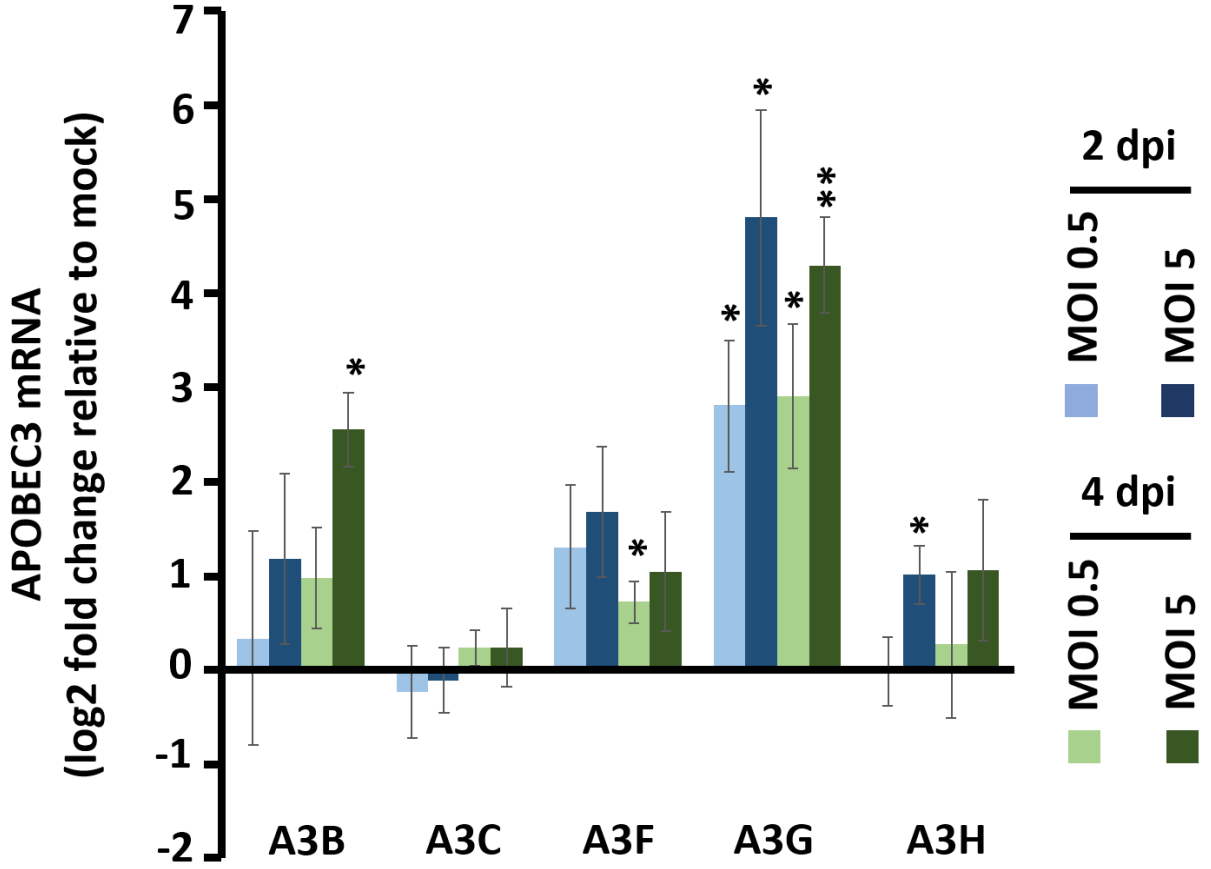

*Figure 24: A3G expression is significantly upregulated upon SARS-CoV-2 infection of HBEC3-KT-ACE2 cells. Bronchial cells have been infected with SARS-CoV-2 at an MOI of 0,5 and 5. Cells have been harvested 2- and 4-days post-infection. Cellular RNA has been extracted and the level of the A3s mRNA was assessed by RTqPCR. Ct between non- and infected cells are compared and the fold change is plotted. Comparison was done by Student's t-test, *indicates P-value < 0,05, **indicates P-value < 0,01.*

HBEC3-KT-ACE2 cells infected with SARS-CoV-2 express A3G proteins

Based on the increased level of A3G RNA expression, we chose to monitor A3G induction at the protein level. We wanted to know whether the increase in A3G RNA expression was also reflected at the protein level. Thus, following the initial infection, we harvested cells and extracted proteins from them. We analyzed the A3G protein level by Western blot using an antibody that detects A3A, A3B and A3G. A3A, A3B and A3G are expected to migrate respectively at 29, 37 and 40kDa. In mock conditions, no A3G proteins are observed while A3G signal can be seen in infected cells. The A3G signal level increases in a dose- and time-dependent manner (Figure 25). We quantified the level of induction by densitometry analysis. We compared the level of expressed proteins in infected cells to that in non-infected cells and plotted the fold change. We used Hsp90 as loading control and used it to normalize the protein signal. No significant fold change is observed at 2 days post-infection. At 4 days post-infection, significant fold change is measured in both infection conditions: MOI of 0,5 and 5 (Figure 26). Protein expression of A3C, A3F and A3H have also been assessed but revealed no protein induction.

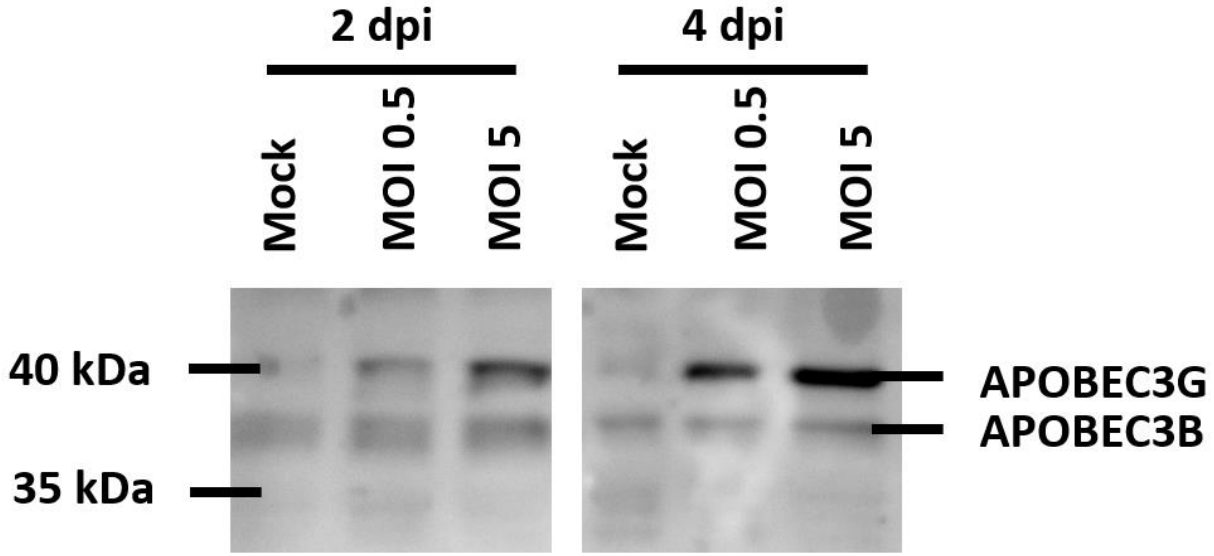

Figure 25: SARS-CoV-2 infection of HBEC3-KT-ACE2 cells induces A3G protein expression. Infected cells have been harvested and protein extracted from them. A3G protein induction is assessed by Western blot.

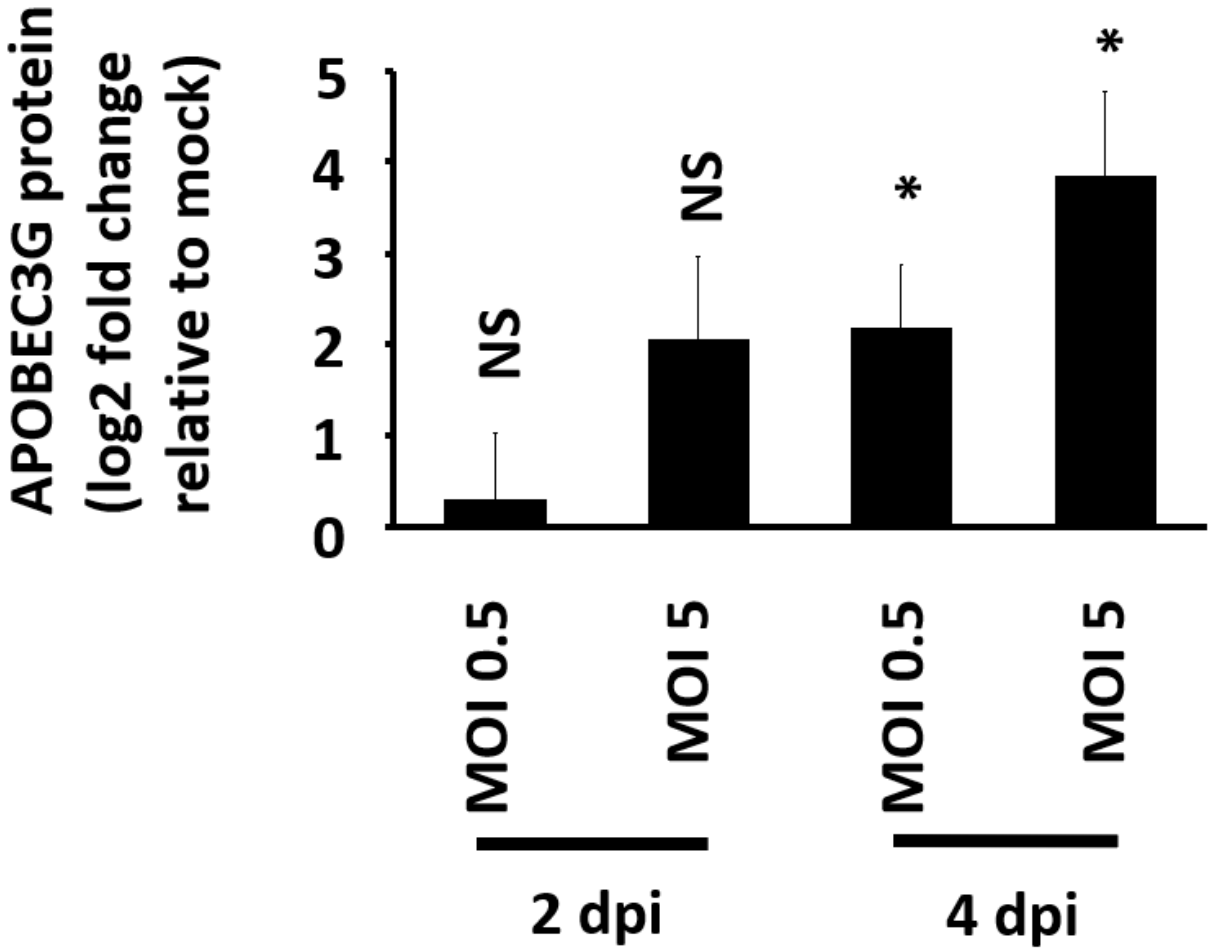

*Figure 26: A3G protein level is significantly increased in SARS-CoV-2 infected HBEC3-KT-ACE2 cells. The level of A3G expression in infected and non-infected cells have been measured and compared by densitometry. Protein signal is normalized to Hsp90 signal level and compared to mock cells. Comparison was done by Student's t-test, NS indicates P-value > 0,05, *indicates P-value < 0,05.*

Deaminase activity is not detected in SARS-CoV-2 infected HBEC3-KT-ACE2 cells

Given the increased expression of A3G both at the mRNA level and the protein level, we chose to assess the deaminase activity from the infected and non-infected cells. For this, we conducted a deamination test. This assay is based on the deamination of a fluorescently-labeled DNA probe bearing the favored deamination motif of A3G: 5'-CC-3' (TCC probe). When deamination activity is present, the DNA probe is deaminated and cytosine is converted into uracil. Upon detection, uracil is excised from the DNA backbone by an added enzyme, the Uracil-DNA glycosylase. The reaction mix is treated with NaOH and heated, which induces the breakage of the probe. The product is loaded in a urea gel and migrated. Given that the probe is only marked on one side, the rupture makes the labeled probe fragment shorter and therefore it migrates further. Simply put, when deamination activity is present, the probe migrates further. The positive control condition consists of a TCC probe and a pre-deaminated TUA probe. This control condition must migrate further as the cytosine is already deaminated. Uracil will therefore be automatically excised. As expected, no deamination activity is observed in mock cells since A3G is not expressed in non-infected cells. A similar signal is observed in infected cells (Figure 27). Densitometry analysis was used for quantification and the positive control was used to normalize signals. No significant fold change is measured (Figure 28).

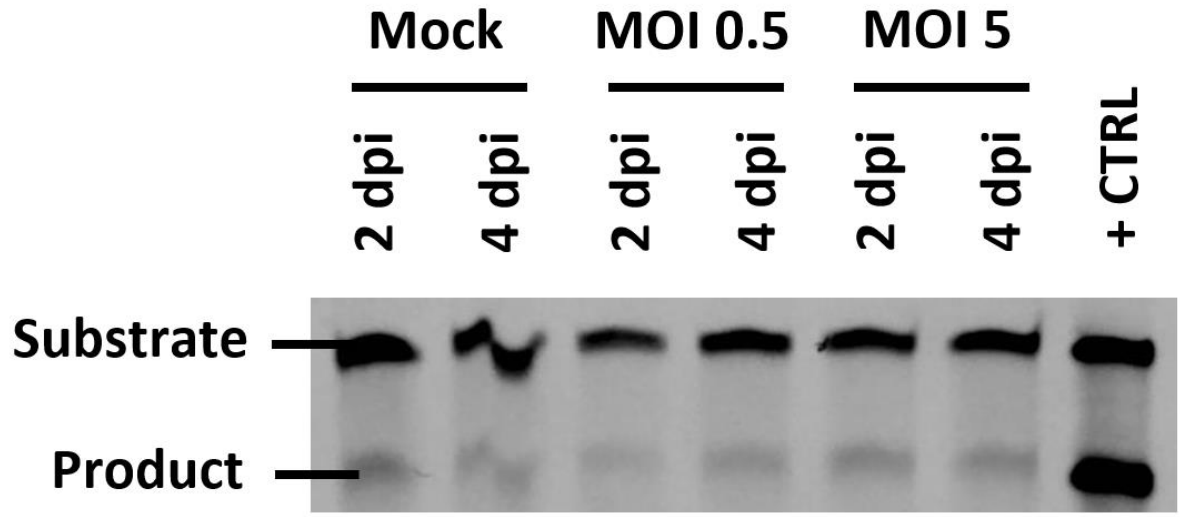

Figure 27: No increase of deaminase activity is observed in infected and A3G expressing HBEC3-KT-ACE2 cells. Infected cells have been harvested and extracted in non-denaturing conditions. Extracted proteins were loaded in a urea gel. Deaminase activity is assessed through the migration of fluorescently-labeled DNA probe bearing A3G favored motif (5'-CC-3') and susceptible to deamination.

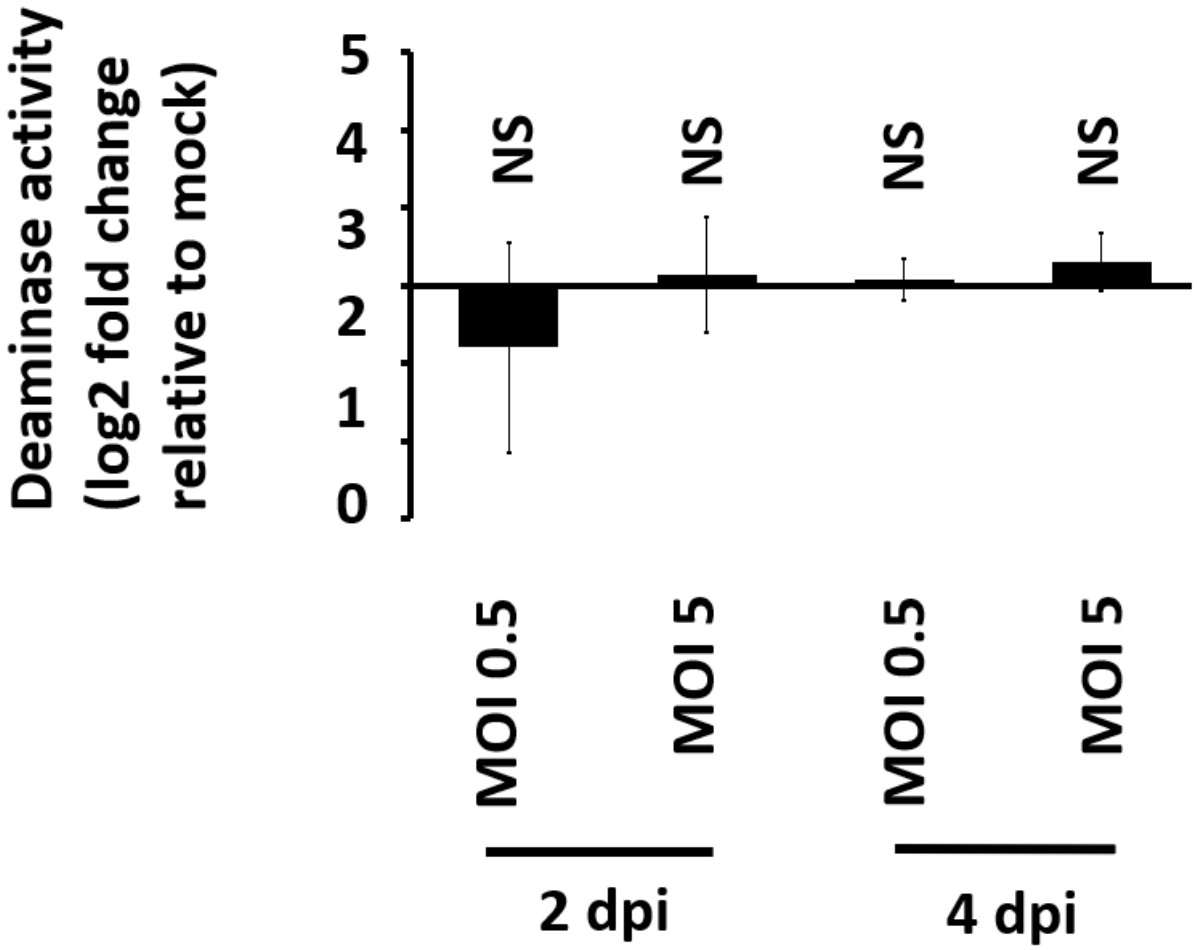

Figure 28: No change of deaminase activity is observed in infected cells compared to mock-infected. Mock cells and infected cells have been compared by densitometry analysis. The positive control is used for normalization. Comparison was done by Student's t-test, NS indicates P-value > 0,05.

Discussion

Throughout the coexistence between pathogens and hosts, a perpetual arms race has been taking place. At every moment the viruses try to take over their hosts. Fortunately, we can defend ourselves thanks to our immune system, which is full of defense mechanisms. Among those are the APOBEC proteins. These proteins are innate immune effectors and restrict mobile genetic elements and viruses. Thus, in 2002, HIV was the first virus shown to be susceptible to APOBEC restriction (Sheehy *et al.,* 2002). Over the years, a multitude of other viruses has also shown their sensitivity to APOBEC restriction. The restriction is based on two types of mechanisms: deaminase-dependent and -independent. When viruses are restricted in a deaminase-dependent way, the footprint of the APOBEC proteins can be observed. Therefore, an APOBEC signature can be detected when analyzing viral genomes. It is found in a multitude of viruses including endemic *Coronaviridae* (Poulain *et al.,* 2020).

Because of its pandemic nature, a lot of raw data has been collected and is now available. Obviously, there are still many unknowns about SARS-CoV-2, but the huge amount of data allows an in-depth study of its genome throughout the evolution of the pandemic. Thus, several authors have decided to track the viral evolution. To do so, they compared a panel of viral genomes collected from infected patients throughout the pandemic with one of the very first

collected. Because of its early origin, it is assumed to be very close to the original ancestor. These analyses revealed that over time, biases in the distribution of nucleotides appeared in the SARS-CoV-2 genome (Mourier *et al.,* 2021; Sadykov *et al.,* 2021; R. Wang *et al.,* 2020). When these biases are clearly defined by an element or an event, they are called mutational signatures (Kockler & Gordenin, 2021). The most strongly found bias is the transition from cytosine into uracil. This mark is one of the characteristics of the APOBEC signature. Moreover, by observing the nucleotide context around the biased areas, it was noticed that this context also corresponds to that favored by the APOBEC proteins (Di Giorgio *et al.,* 2020). Therefore, we can suspect them to be at the origin of the emergence of such nucleotide biases. Just as the APOBEC signature has been observed in human endemic coronaviruses, APOBEC proteins could also be responsible for it in SARS-CoV-2. However, this signature has not been observed in other coronaviruses of zoonotic origin that have also caused epidemics (SARS-CoV and MERS-CoV in 2002 and 2012, respectively) (Poulain *et al.,* 2020). This may be explained by the fact that these viruses replicated in a relatively low number of infected individuals, therefore making the accumulation of mutations not yet sufficient to constitute a detectable evolutionary footprint. The SARS-CoV-2 pandemic is of a completely different magnitude, both geographically and in terms of the number of infected people. In addition to the observation of this mutational footprint that may be attributed to APOBEC proteins, it has recently been shown that certain APOBEC proteins are indeed capable of restricting HCoV-NL63, a human endemic coronavirus (Milewska *et al.,* 2018).

Based on these evidences, we chose to study the impact of the APOBEC3 innate immune effectors on SARS-CoV-2 replication. Thus, we started by investigating the effect of APOBEC proteins on the replication kinetics of the virus. To do so, we designed several Grivet monkey (*Cercopithecus aethiops*) kidney cell lines genetically modified to overexpress the different APOBEC proteins. Unfortunately, due to some design problems, we were not able to measure the impact of all APOBEC proteins. Indeed, we could not detect the presence of A1, A3DE and A3G in the respective cell lineages by western blot. To find the origin of this failure we analyzed the sequence of the transgene. We observed that some cell lines had undergone non-homogenous transduction where the transgene was not always correctly spliced. Therefore, we only measured the impact of A2, A3A, A3B, A3C, A3F and A3H on the SARS-CoV-2 replication, since the other cell lines were not correctly established. For this purpose, we infected each APOBEC-overexpressing cell lines and measured the replication kinetics of SARS-CoV-2 under these different conditions. We were not able to show any obvious difference in cytopathic effect or particular impact on virus replication compared to cell lines not expressing APOBEC proteins.

In order to guide us on which A3 proteins would be the most relevant to study in the context of SARS-CoV-2 replication, we decided to study their expression profile in human bronchial cells infected with SARS-CoV-2. We chose to use the immortalized primary bronchial cells HBEC3-KT and transduced them with an ACE2 expression vector to make them permissive to SARS-CoV-2 infection. We compared infected cells with uninfected cells and it turned out that among A3s, the expression of A3G seems to be particularly upregulated at the mRNA and protein level both in a dose and time-dependent manner. Besides the RNA expression and the protein production, there is a third level to investigate: the activity of the protein. Indeed, for many reasons, it is possible to observe the presence of the protein but not its activity. Thus, APOBEC proteins being deaminases, we tested their ability to deaminate a DNA substrate bearing their preferred motif. Importantly, we could not observe any increased deaminase activity in the infected cells despite an increase in A3G protein level. To sum up: upon SARS-CoV-2 infection, bronchial cells express A3G mRNA, produce A3G proteins but its activity cannot be detected. There are several possible explanations. Firstly, from a technical point of view, we

loaded about 60µg of protein onto our urea gels for the deamination assay. From experience in the lab, we know that this amount can sometimes be too low for the detection of the protein activity. Secondly, there are many biological reasons that may be behind the loss of deaminase activity. It would not be surprising that the virus expresses a viral protein dedicated to the antagonization of antiviral proteins like A3G. In this case, the inhibition of deaminase activity would be a defense mechanism set up by the virus to counteract the action of APOBEC proteins. Such mechanisms are not unknown and have already been observed, for example in EBV. Indeed, its BORF2 protein can stoichiometrically inhibits A3B enzymatic activity (Cheng *et al.*, 2019). This type of mechanism has not been yet observed in *Coronaviridae*, but the question deserves further investigation. On the other hand, a study on the SARS-CoV-2 interactome has suggested a potential escape mechanism to APOBEC response. Indeed, a unique ORF is found within the SARS-CoV-2 genome, ORF10. The 38 amino acid long peptide has been shown to interact with the cullin-2 RING E3 ligase complex. This is strongly reminiscent of the HIV Vif protein that antagonizes A3G by promoting its interaction with the CUL2-ubiquitin ligase complex and its subsequent proteasomal degradation. Through this interaction, ORF10 might hijack the ubiquitination pathway and play a Vif-like role (Gordon *et al.*, 2020). A recent study casts doubt on this hypothesis. In fact, two polish patients were infected by a SARS-CoV-2 variant that was not expressing ORF10. The ORF10 sequence was actually disrupted by a premature stop codon. A C-to-U transition led to the mutation of CAA into TAA. However, both patients developed symptoms and the first infected the second. This variant seems still pathogenic and infectious. Samples from these patients were obtained for in vitro investigation. This variant showed no obvious difference in fitness compared to other variants commonly found in the Polish population (Pancer *et al.*, 2020). In fact, the ORF10 is not the only actor suspected to be involved in the escape of the immune response, other elements guide us to potential leads to follow. For example, A3G has been shown to interact with the N protein of the SARS-CoV-1 (S. M. Wang & Wang, 2009). This interaction allows A3G to be packaged into the virions during the budding. But in the case of HTLV-1, it prevents the encapsidation by crowding out A3G from the budding site (Derse *et al.*, 2007). As seen before, HTLV-1 NC bears a particular motif to inhibit A3G activity. This motif is partially found in SARS-CoV N protein and could allow SARS-CoV to escape the APOBEC response. SARS-CoV and SARS-CoV2 are indeed different viruses but they are very close and their N protein shares 90% homology (Grifoni *et al.*, 2020). The nucleocapsid could therefore be a key factor in the escape of the immune system. In addition, the N protein can antagonize the IFN signaling pathway. By interfering with RIG-I receptor, it represses the IFN-β production and by preventing nuclear translocation of transcription factors, it inhibits the expression of IFN stimulated genes (Bai *et al.*, 2021). APOBEC proteins are precisely induced by IFN (Covino *et al.*, 2018). Multiple SARS-CoV-2 proteins are actually involved in the repression of the interferon pathway, either directly or through its signaling. In fact, a large part of SARS-CoV2 proteins has been shown to modulate IFN pathways to varying degrees. Nsp1 inhibits host translation and thus the synthesis of interferon-induced proteins. Nsp6, nsp13 and ORF6 inhibit IFN-1 expression. Nsp6, nsp13, ORF6, ORF7a, ORF7b interfere with the IFN-1 signaling by inhibiting phosphorylation events in the IFN pathway (Lei *et al.*, 2020; W. Wang *et al.*, 2021; Xia *et al.*, 2020; Yuen *et al.*, 2020). Moreover, when comparing infected patients, those with severe disease show a strong downregulation of antiviral genes such as A3G compared to those with moderate symptoms (S. Li *et al.*, 2021). Also, people with innate deficiency in IFN-1 immunity show poorer clinical outcomes. Furthermore, IFN-β treatment has been shown to effectively inhibit SARS-CoV-2 replication (Lei *et al.*, 2020). These studies highlight the importance of the IFN pathway in the control of SARS-CoV-2 infection and its suppression could be a key mechanism used by the virus to modulate the immune response, including APOBEC proteins.

There is still much mystery surrounding the biology of SARS-CoV-2, the relation between APOBEC proteins and SARS-CoV2 is even more so. We were not able to show an effect on the viral replication from APOBEC proteins, yet we did point the induction of A3G mRNA and protein upon infection. This observation is one of the first steps in the understanding of SARS-CoV-2 biology and how the virus can antagonize the immune system and its effectors. Many questions still need to be answered. How is the induction set up and by what mechanism? Is A3G really the only one induced? We can indeed see a slight upregulation of A3B, A3F and A3H expression. But in these cases, we could not detect the proteins. As in the case of HIV, A3G may be the primary actor in the restriction and the other family members may provide milder effects. We may not be able to detect the other APOBEC proteins due to a degradation mechanism. It is conceivable that the virus could inhibit A3G activity while degrading the other APOBEC proteins. To answer this question, we could measure the kinetics of infection with an ORF10-KO SARS-CoV-2 in APOBEC overexpressing cell lines and compare its fitness to wild-type virus. We could also follow the induction of APOBEC proteins in SARS-CoV-2 infected cells treated with different levels of IFN, as well as the kinetic of infection of SARS-CoV-2 in response to these IFN levels. The virus may indeed be defending itself by inhibiting APOBEC proteins at the source. This also could be done in APOBEC-deficient cell lines to possibly highlight the importance of APOBEC proteins and the effect of their absence. Or on the contrary we could use cells deficient in other antiviral actors to monitor only the effect of APOBEC proteins. As in the case of ORF10, multiple mutants could also be used to monitor the effect of such deletion on APOBEC induction. It could also be interesting to look at the subcellular distribution of the APOBEC proteins in infected and uninfected cells. Indeed, coronaviruses replicate in perinuclear double-membrane structures hiding the replication from the intracellular innate sensors. If co-localization is observed, how does the virus deal with the presence of APOBEC proteins near its replication site? Does the virus have an inhibition factor against our deaminases? If no co-localization is observed, is there an active viral mechanism relocating the deaminases away from the viral replication centers?

Besides these questions and our first responses, we already planned more experiments for the future. Indeed, the investigation of the link between APOBEC proteins and SARS-CoV-2 does not end here. First, even though we are confident that the band observed by Western blot is rightly attributed to A3G, we plan to comfort this attribution by using a modified cell line that can no longer express A3G. We constructed a lentiviral vector expressing a short hairpin RNA (shRNA) designed to inhibit A3G (shA3G). This lentiviral vector will be used to transduce HBEC3-KT-ACE2 cells and to shut down A3G production. This new cell line will be infected by the SARS-CoV-2 and A3G protein expression will be followed by WB. If we can no longer detect A3G proteins in this lineage, we can confirm that it was indeed A3G that we detected in the first infection assay. In parallel, we also constructed a cell line that expresses a scrambled shRNA (shSCr). This shRNA does not bind A3G mRNA and will serve as control. Besides these two shRNA-carrying vectors, we have also reconstructed the A3G over-expressing lentiviral vectors that were used in the Vero E6 infection. Thus, we started the establishment of four new cell lines: HBEC3-KT-ACE2-shA3G/shSCr/A3G/GFP. This will allow us to conduct another assay of replication kinetics. Unlike the first assay we ran, we plan to change some parameters. First, we will infect at a lower MOI, we think that the first MOI was too high to see an APOBEC effect. Secondly, we will use a recombinant SARS-CoV-2 virus encoding the neon green protein (described in (Xie *et al.,* 2020), already in use in the lab). These new conditions will allow us to assess the replication kinetics by RTqPCR both in the supernatant and in the cells, by flow cytometry and by fluorescent microscopy where we will be able to follow the percentage of neon green cells. Cell viability will be assessed by ATP measurement and cytopathic effects will be monitored by microscopy as done before. The virions present in the supernatant will also be titrated by TCID50 assay (50% Tissue Culture Infectious Dose). We

will compare the replication kinetics between cells that do not express A3G (HBEC3-KT-ACE2-ShA3G), cells that overexpress it (HBEC3-KT-ACE2-A3G) and cells that express it naturally (HBEC3-KT-ACE2-ShSCr/GFP). We will also measure the protein expression of A3G and its deaminase activity in the different conditions. Finally, thanks to the fluorescent virus we will also be able to monitor the expression of HA-tagged A3G in the infected (neon green positive) and uninfected (neon green negative) cells. If the level of A3G decreases in the infected cells, this could indicate that the virus can degrade the protein.

All these important questions still need to be answered. Many are still unresolved and deserve to be. Indeed, the COVID-19 outbreak is still ongoing and has caused hundreds of millions of cases and thousands of deaths are listed each day in the world. Finding answers to these questions and identifying viral mechanisms antagonizing innate effectors would be a major contribution, notably for the development of new therapeutic agents. Some of the answers could be key factors for the understanding of virus evolution. Understanding the drivers of viral evolution reveal to be essential to cope with the emergence of new SARS-CoV-2 variants as well as facing potential future viral epidemics.

Bibliography

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V, Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A. L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjörd, J. E., Foekens, J. A., … Stratton, M. R. (2013). Signatures of mutational processes in human cancer. *Nature*, *500*(7463), 415–421. https://doi.org/10.1038/nature12477

Anwar, F., Davenport, M. P., & Ebrahimi, D. (2013). Footprint of APOBEC3 on the Genome of Human Retroelements. *Journal of Virology*, *87*(14), 8195–8204. https://doi.org/10.1128/JVI.00298-13/ASSET/FF70C3F7-612A-4A97-8E94-3BB5E5D3AE1F/ASSETS/GRAPHIC/ZJV9990978640009.JPEG

Arakawa, H., HauschiLd, J., & Buerstedde, J. M. (2002). Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science (New York, N.Y.)*, *295*(5558), 1301–1306. https://doi.org/10.1126/SCIENCE.1067308

Baggen, J., Vanstreels, E., Jansen, S., & Daelemans, D. (2021). Cellular host factors for SARS-CoV-2 infection. *Nature Microbiology*, *6*(10), 1219–1232. https://doi.org/10.1038/s41564-021-00958-0

Bai, Z., Cao, Y., Liu, W., Li, J., Lundstrom, K., & Davidson, A. (2021). *The SARS-CoV-2 Nucleocapsid Protein and Its Role in Viral Structure, Biological Functions, and a Potential Target for Drug or Vaccine Mitigation*. https://doi.org/10.3390/v13061115

Bekerman, E., Jeon, D., Ardolino, M., & Coscoy, L. (2013). A role for host activation-induced cytidine deaminase in innate immune defense against KSHV. *PLoS Pathogens*, *9*(11). https://doi.org/10.1371/JOURNAL.PPAT.1003748

Bennett, R. P., Diner, E., Sowden, M. P., Lees, J. A., Wedekind, J. E., & Smith, H. C. (2006). APOBEC-1 and AID are nucleo-cytoplasmic trafficking proteins but APOBEC3G cannot traffic. *Biochemical and Biophysical Research Communications*, *350*(1), 214–219. https://doi.org/10.1016/J.BBRC.2006.09.032

Bennett, R. P., Presnyak, V., Wedekind, J. E., & Smith, H. C. (2008). Nuclear Exclusion of the HIV-1 host defense factor APOBEC3G requires a novel cytoplasmic retention signal and is not dependent on RNA binding. *The Journal of Biological Chemistry*, *283*(12), 7320–7327. https://doi.org/10.1074/JBC.M708567200

Biémont, C., & Vieira, C. (2006). Junk DNA as an evolutionary force. *Nature 2006 443:7111*, *443*(7111), 521–524. https://doi.org/10.1038/443521a

Blanc, V., Henderson, J. O., Newberry, E. P., Kennedy, S., Luo, J., & Davidson, N. O. (2005). Targeted deletion of the murine apobec-1 complementation factor (acf) gene results in embryonic lethality. *Molecular and Cellular Biology*, *25*(16), 7260–7269. https://doi.org/10.1128/MCB.25.16.7260-7269.2005

Bogerd, H. P., Doehle, B. P., Wiegand, H. L., & Cullen, B. R. (2004). A single amino acid difference in the host APOBEC3G protein controls the primate species specificity of HIV type 1 virion infectivity factor. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(11), 3770–3774. https://doi.org/10.1073/PNAS.0307713101

Boson, B., Legros, V., Zhou, B., Siret, E., Mathieu, C., Cosset, F. L., Lavillette, D., & Denolly, S. (2021). The SARS-CoV-2 envelope and membrane proteins modulate

maturation and retention of the spike protein, allowing assembly of virus-like particles. *The Journal of Biological Chemistry*, *296*. https://doi.org/10.1074/JBC.RA120.016175

Burns, M. B., Leonard, B., & Harris, R. S. (2015). APOBEC3B: Pathological consequences of an innate immune DNA mutator. *Undefined*, *38*(2), 102–110. https://doi.org/10.4103/2319-4170.148904

Chan, J. F. W., Kok, K. H., Zhu, Z., Chu, H., To, K. K. W., Yuan, S., & Yuen, K. Y. (2020). Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes & Infections*, *9*(1), 221. https://doi.org/10.1080/22221751.2020.1719902

Chen, H., Lilley, C. E., Yu, Q., Lee, D. V., Chou, J., Narvaiza, I., Landau, N. R., & Weitzman, M. D. (2006). APOBEC3A Is a Potent Inhibitor of Adeno-Associated Virus and Retrotransposons. *Current Biology*, *16*(5), 480–485. https://doi.org/10.1016/J.CUB.2006.01.031

Chen, S. H., Habib, G., Yang, C. Y., Gu, Z. W., Lee, B. R., Weng, S. A., Silberman, S. R., Cai, S. J., Deslypere, J. P., Rosseneu, M., Gotto, A. M., Li, W. H., & Chan, L. (1987). Apolipoprotein B-48 Is the Product of a Messenger RNA with an Organ-Specific In-Frame Stop Codon. *Science*, *238*(4825), 363–366. https://doi.org/10.1126/SCIENCE.3659919

Chen, Y., Liu, Q., & Guo, D. (2020). Emerging coronaviruses: Genome structure, replication, and pathogenesis. *Journal of Medical Virology*, *92*(4), 418. https://doi.org/10.1002/JMV.25681

Cheng, A. Z., Yockteng-Melgar, J., Jarvis, M. C., Malik-Soni, N., Borozan, I., Carpenter, M. A., Mccann, J. L., Ebrahimi, D., Shaban, N. M., Marcon, E., Greenblatt, J., Brown, W. L., Frappier, L., & Harris, R. S. (2019). Epstein-Barr virus BORF2 inhibits cellular APOBEC3B to preserve viral genome integrity. *NAtuRE MICROBIOLOGY* |, *4*, 78–88. https://doi.org/10.1038/s41564-018-0284-6

Chiu, Y. L., & Greene, W. C. (2009). APOBEC3G: an intracellular centurion. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1517), 689. https://doi.org/10.1098/RSTB.2008.0193

Chiu, Y. L., Witkowska, H. E., Hall, S. C., Santiago, M., Soros, V. B., Esnault, C., Heidmann, T., & Greene, W. C. (2006). High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proceedings of the National Academy of Sciences*, *103*(42), 15588–15593. https://doi.org/10.1073/PNAS.0604524103

Conticello, S. G., Thomas, C. J. F., Petersen-Mahrt, S. K., & Neuberger, M. S. (2005). Evolution of the AID/APOBEC Family of Polynucleotide (Deoxy)cytidine Deaminases. *Molecular Biology and Evolution*, *22*(2), 367–377. https://doi.org/10.1093/MOLBEV/MSI026

Covino, D. A., Gauzzi, M. C., & Fantuzzi, L. (2018). Understanding the regulation of APOBEC3 expression: Current evidence and much to learn. *Journal of Leukocyte Biology*, *103*(3), 433–444. https://doi.org/10.1002/JLB.2MR0717-310R/FORMAT/PDF

David, S. S., O'Shea, V. L., & Kundu, S. (2007). Base-excision repair of oxidative DNA damage. *Nature 2007 447:7147*, *447*(7147), 941–950. https://doi.org/10.1038/nature05978

Davidson, N. O., Innerarity, T. L., Scott, J., Smith, H., Driscoll, D. M., Teng, B., & Chan, L. (1995). Proposed nomenclature for the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme: APOBEC-1. *RNA*, *1*(1), 3. https://www.ncbi.nlm.nih.gov/sites/ppmc/articles/PMC1369056/

Davidson, N. O., & Shelness, G. S. (2000). APOLIPOPROTEIN B: mRNA editing, lipoprotein assembly, and presecretory degradation. *Annual Review of Nutrition*, *20*, 169–193. https://doi.org/10.1146/ANNUREV.NUTR.20.1.169

De Maio, F., Lo Cascio, E., Babini, G., Sali, M., Della Longa, S., Tilocca, B., Roncada, P., Arcovito, A., Sanguinetti, M., Scambia, G., & Urbani, A. (2020). Improved binding of SARS-CoV-2 Envelope protein to tight junction-associated PALS1 could play a key role in COVID-19 pathogenesis. *Microbes and Infection*, *22*(10), 592–597. https://doi.org/10.1016/J.MICINF.2020.08.006

Delviks-Frankenberry, K. A., Desimmie, B. A., & Pathak, V. K. (2020). Structural Insights into APOBEC3-Mediated Lentiviral Restriction. *Viruses 2020, Vol. 12, Page 587*, *12*(6), 587. https://doi.org/10.3390/V12060587

Derse, D., Hill, S. A., Princler, G., Lloyd, P., & Heidecker, G. (2007). Resistance of human T cell leukemia virus type 1 to APOBEC3G restriction is mediated by elements in nucleocapsid. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(8), 2915–2920. www.pnas.orgcgidoi10.1073pnas.0609444104

Di Giorgio, S., Martignano, F., Torcia, M. G., Mattiuz, G., & Conticello, S. G. (2020). Evidence for host-dependent RNA editing in the transcriptome of SARS-CoV-2. *Science Advances*, *6*(25), 1–9. https://doi.org/10.1126/sciadv.abb5813

Eckerle, L. D., Becker, M. M., Halpin, R. A., Li, K., Venter, E., Lu, X., Scherbakova, S., Graham, R. L., Baric, R. S., Stockwell, T. B., Spiro, D. J., & Denison, M. R. (2010). Infidelity of SARS-CoV Nsp14-Exonuclease Mutant Virus Replication Is Revealed by Complete Genome Sequencing. *PLOS Pathogens*, *6*(5), e1000896. https://doi.org/10.1371/JOURNAL.PPAT.1000896

Esnault, C., Millet, J., Schwartz, O., & Heidmann, T. (2006). Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses. *Nucleic Acids Research*, *34*(5), 1522–1531. https://doi.org/10.1093/NAR/GKL054

Estola, T. (1970). Coronaviruses, a new group of animal RNA viruses. *Avian Diseases*, *14*(2), 330–336. https://doi.org/10.2307/1588476

Fehrholz, M., Kendl, S., Prifert, C., Weissbrich, B., Lemon, K., Rennick, L., Duprex, P. W., Rima, B. K., Koning, F. A., Holmes, R. K., Malim, M. H., & Schneider-Schaulies, J. (2012). The innate antiviral factor APOBEC3G targets replication of measles, mumps and respiratory syncytial viruses. *Journal of General Virology*, *93*(3), 565–576. https://doi.org/10.1099/VIR.0.038919-0/CITE/REFWORKS

Forbes, S. A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C. G., Ward, S., Dawson, E., Ponting, L., Stefancsik, R., Harsha, B., YinKok, C., Jia, M., Jubb, H., Sondka, Z., Thompson, S., De, T., & Campbell, P. J. (2017). COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Research*, *45*(Database issue), D777. https://doi.org/10.1093/NAR/GKW1121

Fossat, N., Tourle, K., Radziewic, T., Barratt, K., Liebhold, D., Studdert, J. B., Power, M., Jones, V., Loebel, D. A. F., & Tam, P. P. L. (2014). C to U RNA editing mediated by

APOBEC1 requires RNA-binding protein RBM47. *EMBO Reports*, *15*(8), 903–910. https://doi.org/10.15252/EMBR.201438450

Fugmann, S. D., & Schatz, D. G. (2002). One AID to unite them all. *Science*, *295*(5558), 1244–1245. https://doi.org/10.1126/SCIENCE.1070023/ASSET/B4D2DF62-F752-4F0D-9CF0-9D19DEB4F5F3/ASSETS/SCIENCE.1070023.FP.PNG

Gallois-Montbrun, S., Kramer, B., Swanson, C. M., Byers, H., Lynham, S., Ward, M., & Malim, M. H. (2007). Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules. *Journal of Virology*, *81*(5), 2165–2178. https://doi.org/10.1128/JVI.02287-06

Goila-Gaur, R., Khan, M. A., Miyagi, E., Kao, S., Opi, S., Takeuchi, H., & Strebel, K. (2008). HIV-1 Vif promotes the formation of high molecular mass APOBEC3G complexes. *Virology*, *372*(1), 136. https://doi.org/10.1016/J.VIROL.2007.10.017

Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., … Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, *583*(7816), 459–468. https://doi.org/10.1038/s41586-020-2286-9

Greenwell-Wild, T., Vázquez, N., Jin, W., Rangel, Z., Munson, P. J., & Wahl, S. M. (2009). Interleukin-27 inhibition of HIV-1 involves an intermediate induction of type I interferon. *Blood*, *114*(9), 1864–1874. https://doi.org/10.1182/BLOOD-2009-03-211540

Greeve, J., Altkemper, I., Dieterich, J. H., Greten, H., & Windler, E. (1993). Apolipoprotein B mRNA editing in 12 different mammalian species: hepatic expression is reflected in low concentrations of apoB-containing plasma lipoproteins. *Journal of Lipid Research*, *34*(8), 1367–1383. https://doi.org/10.1016/S0022-2275(20)36966-2

Grifoni, A., Sidney, J., Zhang, Y., Scheuermann, R. H., Peters, B., & Sette, A. (2020). A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host & Microbe*, *27*(4), 671. https://doi.org/10.1016/J.CHOM.2020.03.002

Guerrero, S., Libre, C., Batisse, J., Mercenne, G., Richer, D., Laumond, G., Decoville, T., Moog, C., Marquet, R., & Paillart, J. C. (2016). Translational regulation of APOBEC3G mRNA by Vif requires its 5′UTR and contributes to restoring HIV-1 infectivity. *Scientific Reports 2016 6:1*, *6*(1), 1–13. https://doi.org/10.1038/srep39507

Guo, F., Cen, S., Niu, M., Saadatmand, J., & Kleiman, L. (2006). Inhibition of [... formula ...]-Primed Reverse Transcription by Human APOBEC3G during Human Immunodeficiency Virus Type 1 Replication. *Journal of Virology*, *80*(23), 11710. https://doi.org/10.1128/JVI.01038-06

Guo, Y., Dong, L., Qiu, X., Wang, Y., Zhang, B., Liu, H., Yu, Y., Zang, Y., Yang, M., & Huang, Z. (2014). Structural basis for hijacking CBF-β and CUL5 E3 ligase complex by HIV-1 Vif. *Nature*, *505*(7482), 229–233. https://doi.org/10.1038/NATURE12884

Haque, S. K. M., Ashwaq, O., Sarief, A., & Azad John Mohamed, A. K. (2020). A comprehensive review about SARS-CoV-2. *Future Virology*, *15*(9), 625–648. https://doi.org/10.2217/fvl-2020-0124

Harris, R. S., & Dudley, J. P. (2015). APOBECs and virus restriction. *Virology*, *479–480*,

131–145. https://doi.org/10.1016/J.VIROL.2015.03.012

Harris, R. S., Hultquist, J. F., & Evans, D. T. (2012). The restriction factors of human immunodeficiency virus. *The Journal of Biological Chemistry*, *287*(49), 40875–40883. https://doi.org/10.1074/JBC.R112.416925

Harris, R. S., & Liddament, M. T. (2004). Retroviral restriction by APOBEC proteins. *Nature Reviews Immunology*, *4*(11), 868–877. https://doi.org/10.1038/NRI1489

Hasöksüz, M., Kiliç, S., & Saraç, F. (2020). Coronaviruses and sars-cov-2. *Turkish Journal of Medical Sciences*, *50*(SI-1), 549–556. https://doi.org/10.3906/sag-2004-127

Hayward, J. A., Tachedjian, M., Cui, J., Cheng, A. Z., Johnson, A., Baker, M. L., Harris, R. S., Wang, L. F., & Tachedjian, G. (2018). Differential Evolution of Antiretroviral Restriction Factors in Pteropid Bats as Revealed by APOBEC3 Gene Complexity. *Molecular Biology and Evolution*, *35*(7), 1626–1637. https://doi.org/10.1093/MOLBEV/MSY048

Henderson, S., & Fenton, T. (2015). APOBEC3 genes: retroviral restriction factors to cancer drivers. *Trends in Molecular Medicine*, *21*(5), 274–284. https://doi.org/10.1016/J.MOLMED.2015.02.007

Hernandez, M. M., Fahrny, A., Jayaprakash, A., Gers-Huber, G., Dillon-White, M., Audigé, A., Mulder, L. C. F., Sachidanandam, R., Speck, R. F., & Simon, V. (2020). Impact of Suboptimal APOBEC3G Neutralization on the Emergence of HIV Drug Resistance in Humanized Mice. *Journal of Virology*, *94*(5). https://doi.org/10.1128/JVI.01543-19

Hirano, K. I., Young, S. G., Farese, R. V., Ng, J., Sande, E., Warburton, C., Powell-Braxton, L. M., & Davidson, N. O. (1996). Targeted Disruption of the Mouse apobec-1 Gene Abolishes Apolipoprotein B mRNA Editing and Eliminates Apolipoprotein B48. *Journal of Biological Chemistry*, *271*(17), 9887–9890. https://doi.org/10.1074/JBC.271.17.9887

Hoffmann, M., Hofmann-Winkler, H., Pöhlmann, S., Hoffmann, M., Hofmann-Winkler, · H, & Pöhlmann, · S. (2018). Priming Time: How Cellular Proteases Arm Coronavirus Spike Proteins. *Activation of Viruses by Host Proteases*, 71–98. https://doi.org/10.1007/978-3-319-75474-1_4

Hoffmann, M., Kleine-Weber, H., & Pöhlmann, S. (2020). A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Molecular Cell*, *78*(4), 779. https://doi.org/10.1016/J.MOLCEL.2020.04.022

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N. H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, *181*(2), 271. https://doi.org/10.1016/J.CELL.2020.02.052

Hogg, M. (2010). *Characterising new roles for APOBEC4 and ADAR deaminases* [The University of Edinburgh]. https://era.ed.ac.uk/handle/1842/4792

Holtz, C. M., Sadler, H. A., & Mansky, L. M. (2013). APOBEC3G cytosine deamination hotspots are defined by both sequence context and single-stranded DNA secondary structure. *Nucleic Acids Research*, *41*(12), 6139–6148. https://doi.org/10.1093/nar/gkt246

Hsu, E. (2016). Assembly and Expression of Shark Ig Genes. *The Journal of Immunology*,

196(9), 3517–3523. https://doi.org/10.4049/JIMMUNOL.1600164

Hu, B., Guo, H., Zhou, P., & Shi, Z. L. (2021). Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, *19*(3), 141–154. https://doi.org/10.1038/s41579-020-00459-7

Huang, J., Liang, Z., Yang, B., Tian, H., Ma, J., & Zhang, H. (2007). Derepression of microRNA-mediated protein translation inhibition by apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G) and its family members. *The Journal of Biological Chemistry*, *282*(46), 33632–33640. https://doi.org/10.1074/JBC.M705116200

Huthoff, H., Autore, F., Gallois-Montbrun, S., Fraternali, F., & Malim, M. H. (2009). RNA-Dependent Oligomerization of APOBEC3G Is Required for Restriction of HIV-1. *PLOS Pathogens*, *5*(3), e1000330. https://doi.org/10.1371/JOURNAL.PPAT.1000330

ICTV. (2020). *Nidovirales - Positive Sense RNA Viruses - ICTV 9th Report (2011)*. International Committee on Taxonomy of Viruses. https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/219/nidovirales

Institut Pasteur. (2011). *SRAS*. Institut Pasteur. https://www.pasteur.fr/fr/centre-medical/fiches-maladies/sras

Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M., & Honjo, T. (2004). *Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1*. www.pnas.orgcgidoi10.1073pnas.0307335101

Jaguva Vasudevan, A. A., Perković, M., Bulliard, Y., Cichutek, K., Trono, D., Häussinger, D., & Münk, C. (2013). Prototype foamy virus Bet impairs the dimerization and cytosolic solubility of human APOBEC3G. *Journal of Virology*, *87*(16), 9030–9040. https://doi.org/10.1128/JVI.03385-12

Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J., & Navaratnam, N. (2002). An Anthropoid-Specific Locus of Orphan C to U RNA-Editing Enzymes on Chromosome 22. *Genomics*, *79*(3), 285–296. https://doi.org/10.1006/GENO.2002.6718

Jern, P., Russell, R. A., Pathak, V. K., & Coffin, J. M. (2009). Likely Role of APOBEC3G-Mediated G-to-A Mutations in HIV-1 Evolution and Drug Resistance. *PLoS Pathogens*, *5*(4). https://doi.org/10.1371/JOURNAL.PPAT.1000367

Kahn, J. S., & McIntosh, K. (2005). Discussion. *Pediatric Infectious Disease Journal*, *24*(11 SUPPL.). https://doi.org/10.1097/01.INF.0000188166.17324.60

Kim, E. Y., Lorenzo-Redondo, R., Little, S. J., Chung, Y. S., Phalora, P. K., Maljkovic Berry, I., Archer, J., Penugonda, S., Fischer, W., Richman, D. D., Bhattacharya, T., Malim, M. H., & Wolinsky, S. M. (2014). Human APOBEC3 Induced Mutation of Human Immunodeficiency Virus Type-1 Contributes to Adaptation and Evolution in Natural Infection. *PLoS Pathogens*, *10*(7). https://doi.org/10.1371/journal.ppat.1004281

Kobayashi, M., Takaori-Kondo, A., Miyauchi, Y., Iwai, K., & Uchiyama, T. (2005). *Ubiquitination of APOBEC3G by an HIV-1 Vif-Cullin5-Elongin B-Elongin C Complex Is Essential for Vif Function**. https://doi.org/10.1074/jbc.C500082200

Kockler, Z. W., & Gordenin, D. A. (2021). From RNA World to SARS-CoV-2: The Edited

Story of RNA Viral Evolution. *Cells 2021, Vol. 10, Page 1557*, *10*(6), 1557. https://doi.org/10.3390/CELLS10061557

Kreisberg, J. F., Yonemoto, W., & Greene, W. C. (2006). Endogenous factors enhance HIV infection of tissue naive CD4 T cells by stimulating high molecular mass APOBEC3G complex formation. *Journal of Experimental Medicine*, *203*(4), 865–870. https://doi.org/10.1084/JEM.20051856

Krishnan, A., Iyer, L. M., Holland, S. J., Boehm, T., & Aravind, L. (2018). Diversification of AID/APOBEC-like deaminases in metazoa: Multiplicity of clades and widespread roles in immunity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(14), E3201–E3210. https://doi.org/10.1073/PNAS.1720897115/-/DCSUPPLEMENTAL

Lackey, L., Demorest, Z. L., Land, A. M., Hultquist, J. F., Brown, W. L., & Harris, R. S. (2012). APOBEC3B and AID Have Similar Nuclear Import Mechanisms. *Journal of Molecular Biology*, *419*(5), 301–314. https://doi.org/10.1016/J.JMB.2012.03.011

Lalchhandama, K. (2020). The chronicles of coronaviruses: the bronchitis, the hepatitis and the common cold. *Science Vision*, *20*(1), 43–53. https://doi.org/10.33493/SCIVIS.20.01.04

LaRue, R. S., Andrésdóttir, V., Blanchard, Y., Conticello, S. G., Derse, D., Emerman, M., Greene, W. C., Jónsson, S. R., Landau, N. R., Löchelt, M., Malik, H. S., Malim, M. H., Münk, C., O'Brien, S. J., Pathak, V. K., Strebel, K., Wain-Hobson, S., Yu, X.-F., Yuhki, N., & Harris, R. S. (2009). Guidelines for Naming Nonprimate APOBEC3 Genes and Proteins. *Journal of Virology*, *83*(2), 494–497. https://doi.org/10.1128/JVI.01976-08

Law, E. K., Sieuwerts, A. M., Lapara, K., Leonard, B., Starrett, G. J., Molan, A. M., Temiz, N. A., Vogel, R. I., Meijer-Van Gelder, M. E., Sweep, F. C. G. J., Span, P. N., Foekens, J. A., Martens, J. W. M., Yee, D., & Harris, R. S. (2016). The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. *Science Advances*, *2*(10), 1–10. https://doi.org/10.1126/sciadv.1601737

Lei, X., Dong, X., Ma, R., Wang, W., Xiao, X., Tian, Z., Wang, C., Wang, Y., Li, L., Ren, L., Guo, F., Zhao, Z., Zhou, Z., Xiang, Z., & Wang, J. (2020). Activation and evasion of type I interferon responses by SARS-CoV-2. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-17665-9

Li, J., Chen, Y., Li, M., Carpenter, M. A., McDougle, R. M., Luengas, E. M., Macdonald, P. J., Harris, R. S., & Mueller, J. D. (2014). APOBEC3 Multimerization Correlates with HIV-1 Packaging and Restriction Activity in Living Cells. *Journal of Molecular Biology*, *426*(6), 1296–1307. https://doi.org/10.1016/J.JMB.2013.12.014

Li, S., Duan, X., Li, Y., Li, M., Gao, Y., Li, T., Li, S., Tan, L., Shao, T., Jeyarajan, A. J., Chen, L., Han, M., Lin, W., & Li, X. (2021). Differentially expressed immune response genes in COVID-19 patients based on disease severity. *Aging*, *13*(7), 9265–9276. https://doi.org/10.18632/AGING.202877

Li, Z., Ning, S., Su, X., Liu, X., Wang, H., Liu, Y., Zheng, W., Zheng, B., Yu, X. F., & Zhang, W. (2018). Enterovirus 71 antagonizes the inhibition of the host intrinsic antiviral factor A3G. *Nucleic Acids Research*, *46*(21), 11514–11527. https://doi.org/10.1093/NAR/GKY840

Liao, W., Hong, S. H., Chan, B. H. J., Rudolph, F. B., Clark, S. C., & Chan, L. (1999).

APOBEC-2, a Cardiac- and Skeletal Muscle-Specific Member of the Cytidine Deaminase Supergene Family. *Biochemical and Biophysical Research Communications*, *260*(2), 398–404. https://doi.org/10.1006/BBRC.1999.0925

Liu, M. C., Liao, W. Y., Buckley, K. M., Yang, S. Y., Rast, J. P., & Fugmann, S. D. (2018). AID/APOBEC-like cytidine deaminases are ancient innate immune mediators in invertebrates. *Nature Communications*, *9*(1). https://doi.org/10.1038/S41467-018-04273-X

Löchelt, M., Romen, F., Bastone, P., Muckenfuss, H., Kirchner, N., Kim, Y. B., Truyen, U., Rösler, U., Battenberg, M., Saib, A., Flory, E., Cichutek, K., & Münk, C. (2005). The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(22), 7982. https://doi.org/10.1073/PNAS.0501445102

Luo, K., Wang, T., Liu, B., Tian, C., Xiao, Z., Kappes, J., & Yu, X.-F. (2007). Cytidine Deaminases APOBEC3G and APOBEC3F Interact with Human Immunodeficiency Virus Type 1 Integrase and Inhibit Proviral DNA Formation. *Journal of Virology*, *81*(13), 7238. https://doi.org/10.1128/JVI.02584-06

Malim, M. H., & Pollpeter, D. (2018). APOBEC restriction goes nuclear. *Nature Microbiology 2018 4:1*, *4*(1), 6–7. https://doi.org/10.1038/s41564-018-0323-3

Malone, B., Urakova, N., Snijder, E. J., & Campbell, E. A. (2021). Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. *Nature Reviews Molecular Cell Biology 2021*, 1–19. https://doi.org/10.1038/s41580-021-00432-z

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., & Trono, D. (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature*, *424*(6944), 99–103. https://doi.org/10.1038/NATURE01709

Mangeat, B., Turelli, P., Liao, S., & Trono, D. (2004). A Single Amino Acid Determinant Governs the Species-specific Sensitivity of APOBEC3G to Vif Action. *Journal of Biological Chemistry*, *279*(15), 14481–14483. https://doi.org/10.1074/JBC.C400060200

Mariani, R., Chen, D., Schröfelbauer, B., Navarro, F., König, R., Bollman, B., Münk, C., Nymark-McMahon, H., & Landau, N. R. (2003). Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell*, *114*(1), 21–31. https://doi.org/10.1016/S0092-8674(03)00515-4

Martinez, T., Shapiro, M., Bhaduri-McIntosh, S., & MacCarthy, T. (2019). Evolutionary effects of the AID/APOBEC family of mutagenic enzymes on human gamma-herpesviruses. *Virus Evolution*, *5*(1). https://doi.org/10.1093/VE/VEY040

Matyášek, R., & Kovařík, A. (2020). Mutation Patterns of Human SARS-CoV-2 and Bat RaTG13 Coronavirus Genomes Are Strongly Biased Towards C>U Transitions, Indicating Rapid Evolution in Their Hosts. *Genes 2020, Vol. 11, Page 761*, *11*(7), 761. https://doi.org/10.3390/GENES11070761

Mbisa, J. L., Barr, R., Thomas, J. A., Vandegraaff, N., Dorweiler, I. J., Svarovskaia, E. S., Brown, W. L., Mansky, L. M., Gorelick, R. J., Harris, R. S., Engelman, A., & Pathak, V. K. (2007). Human Immunodeficiency Virus Type 1 cDNAs Produced in the Presence of APOBEC3G Exhibit Defects in Plus-Strand DNA Transfer and Integration. *Journal of Virology*, *81*(13), 7099. https://doi.org/10.1128/JVI.00272-07

Mbisa, J. L., Bu, W., & Pathak, V. K. (2010). APOBEC3F and APOBEC3G Inhibit HIV-1 DNA Integration by Different Mechanisms. *Journal of Virology*, *84*(10), 5250–5259. https://doi.org/10.1128/JVI.02358-09/ASSET/F62B2509-7F8B-4EA6-928D-E330E8D84FD5/ASSETS/GRAPHIC/ZJV9990932300005.JPEG

McDaniel, Y. Z., Wang, D., Love, R. P., Adolph, M. B., Mohammadzadeh, N., Chelico, L., & Mansky, L. M. (2020). Deamination hotspots among APOBEC3 family members are defined by both target site sequence context and ssDNA secondary structure. *Nucleic Acids Research*, *48*(3), 1353–1371. https://doi.org/10.1093/NAR/GKZ1164

McDougall, W. M., & Smith, H. C. (2011). Direct evidence that RNA inhibits APOBEC3G ssDNA cytidine deaminase activity. *Biochemical and Biophysical Research Communications*, *412*(4), 612–617. https://doi.org/10.1016/J.BBRC.2011.08.009

Meshcheryakova, A., Pietschmann, P., Zimmermann, P., Rogozin, I. B., & Mechtcheriakova, D. (2021). AID and APOBECs as Multifaceted Intrinsic Virus-Restricting Factors: Emerging Concepts in the Light of COVID-19. *Frontiers in Immunology*, *12*(July), 1–8. https://doi.org/10.3389/fimmu.2021.690416

Milewska, A., Kindler, E., Vkovski, P., Zeglen, S., Ochman, M., Thiel, V., Rajfur, Z., & Pyrc, K. (2018). APOBEC3-mediated restriction of RNA virus replication. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-24448-2

Mourier, T., Sadykov, M., Carr, M. J., Gonzalez, G., Hall, W. W., & Pain, A. (2021). Host-directed editing of the SARS-CoV-2 genome. *Biochemical and Biophysical Research Communications*, *538*, 35–39. https://doi.org/10.1016/j.bbrc.2020.10.092

Mukhopadhyay, D., Anant, S., Lee, R. M., Kennedy, S., Viskochil, D., & Davidson, N. O. (2002). C→U Editing of Neurofibromatosis 1 mRNA Occurs in Tumors That Express Both the Type II Transcript and apobec-1, the Catalytic Subunit of the Apolipoprotein B mRNA–Editing Enzyme. *American Journal of Human Genetics*, *70*(1), 38. https://doi.org/10.1086/337952

Münk, C., Willemsen, A., & Bravo, I. G. (2012). An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. *BMC Evolutionary Biology*, *12*(1). https://doi.org/10.1186/1471-2148-12-71

Muto, T., Muramatsu, M., Taniwaki, M., Kinoshita, K., & Honjo, T. (2000). Isolation, tissue distribution, and chromosomal localization of the human activation-induced cytidine deaminase (AID) gene. *Genomics*, *68*(1), 85–88. https://doi.org/10.1006/GENO.2000.6268

Narvaiza, I., Linfesty, D. C., Greener, B. N., Hakata, Y., Pintel, D. J., Logue, E., Landau, N. R., & Weitzman, M. D. (2009). Deaminase-independent inhibition of parvoviruses by the APOBEC3A cytidine deaminase. *PLoS Pathogens*, *5*(5). https://doi.org/10.1371/journal.ppat.1000439

Neogi, U., Shet, A., Sahoo, P. N., Bontell, I., Ekstrand, M. L., Banerjea, A. C., & Sonnerborg, A. (2013). Human APOBEC3G-mediated hypermutation is associated with antiretroviral therapy failure in HIV-1 subtype C-infected individuals. *Journal of the International AIDS Society*, *16*(1), 10.7448/IAS.16.1.18472. https://doi.org/10.7448/IAS.16.1.18472

Neuman, B. W., Kiss, G., Kunding, A. H., Bhella, D., Baksh, M. F., Connelly, S., Droese, B., Klaus, J. P., Makino, S., Sawicki, S. G., Siddell, S. G., Stamou, D. G., Wilson, I. A., Kuhn, P., & Buchmeier, M. J. (2011). A structural analysis of M protein in coronavirus

assembly and morphology. *Journal of Structural Biology*, *174*(1), 11. https://doi.org/10.1016/J.JSB.2010.11.021

Nieto-Torres, J. L., DeDiego, M. L., Verdiá-Báguena, C., Jimenez-Guardeño, J. M., Regla-Nava, J. A., Fernandez-Delgado, R., Castaño-Rodriguez, C., Alcaraz, A., Torres, J., Aguilella, V. M., & Enjuanes, L. (2014). Severe Acute Respiratory Syndrome Coronavirus Envelope Protein Ion Channel Activity Promotes Virus Fitness and Pathogenesis. *PLoS Pathogens*, *10*(5), 1004077. https://doi.org/10.1371/JOURNAL.PPAT.1004077

OhAinle, M., Kerns, J. A., Malik, H. S., & Emerman, M. (2006). Adaptive Evolution and Antiviral Activity of the Conserved Mammalian Cytidine Deaminase APOBEC3H . *Journal of Virology*, *80*(8), 3853–3862. https://doi.org/10.1128/JVI.80.8.3853-3862.2006

Olson, M. E., Harris, R. S., & Harki, D. A. (2018). APOBEC Enzymes as Targets for Virus and Cancer Therapy. *Cell Chemical Biology*, *25*(1), 36–49. https://doi.org/10.1016/J.CHEMBIOL.2017.10.007

Pancer, K., Milewska, A., Owczarek, K., Dabrowska, A., Kowalski, M., Labaj, P., Branicki, W., Sanak, M., & Pyrc, K. (2020). The SARS-CoV-2 ORF10 is not essential in vitro or in vivo in humans. *PLOS Pathogens*, *16*(12), e1008959. https://doi.org/10.1371/JOURNAL.PPAT.1008959

Papavasiliou, F. N., & Schatz, D. G. (2002). Somatic Hypermutation of Immunoglobulin Genes: Merging Mechanisms for Genetic Diversity. *Cell*, *109*(2), S35–S44. https://doi.org/10.1016/S0092-8674(02)00706-7

Patton, J. R., Hunt, M. D., & Jamieson, A. (n.d.). *Regulation of Gene Expression - Medical Biochemistry*. DoctorLib. Retrieved November 26, 2021, from https://doctorlib.info/medical/biochemistry/37.html

Perelygina, L., Chen, M. hsin, Suppiah, S., Adebayo, A., Abernathy, E., Dorsey, M., Bercovitch, L., Paris, K., White, K. P., Krol, A., Dhossche, J., Torshin, I. Y., Saini, N., Klimczak, L. J., Gordenin, D. A., Zharkikh, A., Plotkin, S., Sullivan, K. E., & Icenogle, J. (2019). Infectious vaccine-derived rubella viruses emerge, persist, and evolve in cutaneous granulomas of children with primary immunodeficiencies. *PLOS Pathogens*, *15*(10), e1008080. https://doi.org/10.1371/JOURNAL.PPAT.1008080

Pham, P., Bransteitter, R., Petruska, J., & Goodman, M. F. (2003). Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature*, *424*(6944), 103–107. https://doi.org/10.1038/NATURE01760

Poulain, F., Lejeune, N., Willemart, K., & Gillet, N. A. (2020). Footprint of the host restriction factors APOBEC3 on the genome of human viruses. *PLoS Pathogens*, *16*(8), 1–30. https://doi.org/10.1371/JOURNAL.PPAT.1008718

Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J., & Scott, J. (1987). A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell*, *50*(6), 831–840. https://doi.org/10.1016/0092-8674(87)90510-1

Prochnow, C., Bransteitter, R., Klein, M. G., Goodman, M. F., & Chen, X. S. (2006). The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature 2006 445:7126*, *445*(7126), 447–451. https://doi.org/10.1038/nature05492

Ratcliff, J., & Simmonds, P. (2021). Potential APOBEC-mediated RNA editing of the

genomes of SARS-CoV-2 and other coronaviruses and its impact on their longer term evolution. *Virology*, *556*(December 2020), 62–72. https://doi.org/10.1016/j.virol.2020.12.018

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Lagelouse, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A. G., Brousse, N., Muramatsu, M., Notarangelo, L. D., Kinoshita, K., Honjo, T., … Durandy, A. (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*, *102*(5), 565–575. https://doi.org/10.1016/S0092-8674(00)00079-9

Richardson, S. R., Narvaiza, I., Planegger, R. A., Weitzman, M. D., & Moran, J. V. (2014). APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. *ELife*, *3*(3). https://doi.org/10.7554/ELIFE.02008

Rogozin, I. B., Basu, M. K., Jordan, I. K., Pavlov, Y. I., & Koonin, E. V. (2005). APOBEC4, a New Member of the AID/APOBEC Family of Polynucleotide (Deoxy)Cytidine Deaminases Predicted by Computational Analysis. *Http://Dx.Doi.Org/10.4161/Cc.4.9.1994*, *4*(9), 1281–1285. https://doi.org/10.4161/CC.4.9.1994

Rogozin, I. B., Iyer, L. M., Liang, L., Glazko, G. V., Liston, V. G., Pavlov, Y. I., Aravind, L., & Pancer, Z. (2007). Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nature Immunology 2007 8:6*, *8*(6), 647–656. https://doi.org/10.1038/ni1463

Sadykov, M., Mourier, T., Guan, Q., & Pain, A. (2021). Short sequence motif dynamics in the SARS-CoV-2 genome suggest a role for cytosine deamination in CpG reduction. *Journal of Molecular Cell Biology*. https://doi.org/10.1093/jmcb/mjab011

Salamango, D. J., Mccann, J. L., Demir, Ö., Brown, W. L., Amaro, R. E., & Harris, R. S. (2018). APOBEC3B Nuclear Localization Requires Two Distinct N-Terminal Domain Surfaces HHS Public Access. *J Mol Biol*, *430*(17), 2695–2708. https://doi.org/10.1016/j.jmb.2018.04.044

Salter, J. D., Bennett, R. P., & Smith, H. C. (2016). The APOBEC Protein Family: United by Structure, Divergent in Function. *Trends in Biochemical Sciences*, *41*(7), 578–594. https://doi.org/10.1016/j.tibs.2016.05.001

Samuel, C. E. (2011). Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. *Virology*, *411*(2), 180–193. https://doi.org/10.1016/j.virol.2010.12.004

Sato, Y., Ohtsubo, H., Nihei, N., Kaneko, T., Sato, Y., Adachi, S. I., Kondo, S., Nakamura, M., Mizunoya, W., Iida, H., Tatsumi, R., Rada, C., & Yoshizawa, F. (2018). Apobec2 deficiency causes mitochondrial defects and mitophagy in skeletal muscle. *The FASEB Journal*, *32*(3), 1428. https://doi.org/10.1096/FJ.201700493R

Sato, Y., Probst, H. C., Tatsumi, R., Ikeuchi, Y., Neuberger, M. S., & Rada, C. (2010). Deficiency in APOBEC2 Leads to a Shift in Muscle Fiber Type, Diminished Body Mass, and Myopathy. *The Journal of Biological Chemistry*, *285*(10), 7111. https://doi.org/10.1074/JBC.M109.052977

Sawyer, S. L., Emerman, M., & Malik, H. S. (2004). Ancient Adaptive Evolution of the Primate Antiviral DNA-Editing Enzyme APOBEC3G. *PLoS Biology*, *2*(9). https://doi.org/10.1371/JOURNAL.PBIO.0020275

Schumann, G. G. (2007). APOBEC3 proteins: major players in intracellular defence against LINE-1-mediated retrotransposition. *Biochemical Society Transactions*, *35*(Pt 3), 637–642. https://doi.org/10.1042/BST0350637

Severe acute respiratory syndrome coronavirus 2 | SWISS-MODEL Repository. (2021). https://swissmodel.expasy.org/repository/species/2697049

Sharma, S., Patnaik, S. K., Thomas Taggart, R., Kannisto, E. D., Enriquez, S. M., Gollnick, P., & Baysal, B. E. (2015). APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. *Nature Communications 2015 6:1*, *6*(1), 1–15. https://doi.org/10.1038/ncomms7881

Sharma, S., Wang, J., Alqassim, E., Portwood, S., Cortes Gomez, E., Maguire, O., Basse, P. H., Wang, E. S., Segal, B. H., & Baysal, B. E. (2019). Mitochondrial hypoxic stress induces widespread RNA editing by APOBEC3G in natural killer cells. *Genome Biology*, *20*(1), 1–17. https://doi.org/10.1186/S13059-019-1651-1/FIGURES/7

Sheehy, A. M., Gaddis, N. C., Choi, J. D., & Malim, M. H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, *418*(6898), 646–650. https://doi.org/10.1038/NATURE00939

Simmonds, P. (2020). Rampant C→U Hypermutation in the Genomes of SARS-CoV-2 and Other Coronaviruses: Causes and Consequences for Their Short- and Long-Term Evolutionary Trajectories. *MSphere*, *5*(3), 1–13. https://doi.org/10.1128/msphere.00408-20

Skuse, G. R., Cappione, A. J., Sowden, M., Metheny, L. J., & Smith, H. C. (1996). The neurofibromatosis type I messenger RNA undergoes base-modification RNA editing. *Nucleic Acids Research*, *24*(3), 478–486. https://doi.org/10.1093/NAR/24.3.478

Smith, H. C., Bennett, R. P., Kizilyer, A., McDougall, W. M., & Prohaska, K. M. (2012). Functions and regulation of the APOBEC family of proteins. *Seminars in Cell & Developmental Biology*, *23*(3), 258–268. https://doi.org/10.1016/J.SEMCDB.2011.10.004

Stenglein, M. D., & Harris, R. S. (2006). *APOBEC3B and APOBEC3F Inhibit L1 Retrotransposition by a DNA Deamination-independent Mechanism* * □ *S*. https://doi.org/10.1074/jbc.M602367200

Stopak, K. S., Chiu, Y. L., Kropp, J., Grant, R. M., & Greene, W. C. (2006). Distinct patterns of cytokine regulation of APOBEC3G expression and activity in primary lymphocytes, macrophages, and dendritic cells. *The Journal of Biological Chemistry*, *282*(6), 3539–3546. https://doi.org/10.1074/JBC.M610138200

Swanton, C., Mcgranahan, N., Starrett, G. J., & Harris, R. S. (2015). *APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity THE IMPORTANCE OF CANCER DIVERSITY*. https://doi.org/10.1158/2159-8290.CD-15-0344

Teng, B. B., Burant, C. F., & Davidson, N. O. (1993). Molecular Cloning of an Apolipoprotein B Messenger RNA Editing Protein. *Science*, *260*(5115), 1816–1819. https://doi.org/10.1126/SCIENCE.8511591

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., & Thiel, V. (2021). Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology*, *19*(3), 155–170. https://doi.org/10.1038/s41579-020-00468-6

Verhalen, B., Starrett, G. J., Harris, R. S., & Jiang, M. (2016). Functional Upregulation of the DNA Cytosine Deaminase APOBEC3B by Polyomaviruses. *Journal of Virology*, *90*(14), 6379–6386. https://doi.org/10.1128/jvi.00771-16

Virology: Coronaviruses. (1968). *Nature 1968 220:5168*, *220*(5168), 650–650. https://doi.org/10.1038/220650b0

Wang, R., Hozumi, Y., Zheng, Y. H., Yin, C., & Wei, G. W. (2020). Host immune response driving SARS-CoV-2 evolution. *Viruses*, *12*(10), 1–20. https://doi.org/10.3390/v12101095

Wang, S. M., & Wang, C. T. (2009). APOBEC3G cytidine deaminase association with coronavirus nucleocapsid protein. *Virology*, *388*(1), 112–120. https://doi.org/10.1016/J.VIROL.2009.03.010

Wang, W., Zhou, Z., Xiao, X., Tian, Z., Dong, X., Wang, C., Li, L., Ren, L., Lei, X., Xiang, Z., & Wang, J. (2021). SARS-CoV-2 nsp12 attenuates type I interferon production by inhibiting IRF3 nuclear translocation. *Cellular & Molecular Immunology 2021 18:4*, *18*(4), 945–953. https://doi.org/10.1038/s41423-020-00619-y

Warren, C. J., Van Doorslaer, K., Pandey, A., Espinosa, J. M., & Pyeon, D. (2015). Role of the host restriction factor APOBEC3 on papillomavirus evolution. *Virus Evolution*, *1*(1), 15. https://doi.org/10.1093/ve/vev015

WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. (2021, November 22). https://covid19.who.int/

Wichroski, M. J., Robb, G. B., & Rana, T. M. (2006). Human Retroviral Host Restriction Factors APOBEC3G and APOBEC3F Localize to mRNA Processing Bodies. *PLOS Pathogens*, *2*(5), e41. https://doi.org/10.1371/JOURNAL.PPAT.0020041

World Health Organisation. (2019). *WHO EMRO | MERS outbreaks | MERS-CoV | Health topics*. World Health Organisation. http://www.emro.who.int/health-topics/mers-cov/mers-outbreaks.html

Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., Hu, Y., Tao, Z. W., Tian, J. H., Pei, Y. Y., Yuan, M. L., Zhang, Y. L., Dai, F. H., Liu, Y., Wang, Q. M., Zheng, J. J., Xu, L., Holmes, E. C., & Zhang, Y. Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, *579*(7798), 265–269. https://doi.org/10.1038/S41586-020-2008-3

Xia, H., Cao, Z., Xie, X., Menachery, V. D., Rajsbaum, R., & Shi, P.-Y. (2020). Evasion of Type I Interferon by SARS-CoV-2 Graphical Abstract Highlights d SARS-CoV-2 proteins antagonize IFN-I production and signaling d Different SARS-CoV-2 proteins inhibit IFN-I response through distinct mechanisms d SARS-CoV, SARS-CoV-2, and MERS-Co. *CellReports*, *33*, 108234. https://doi.org/10.1016/j.celrep.2020.108234

Xie, X., Muruato, A., Lokugamage, K. G., Narayanan, K., Zhang, X., Zou, J., Liu, J., Schindewolf, C., Bopp, N. E., Aguilar, P. V., Plante, K. S., Weaver, S. C., Makino, S., LeDuc, J. W., Menachery, V. D., & Shi, P. Y. (2020). An Infectious cDNA Clone of SARS-CoV-2. *Cell Host and Microbe*, *27*(5), 841-848.e3. https://doi.org/10.1016/J.CHOM.2020.04.004

Yamanaka, S., Balestra, M. E., Ferrell, L. D., Fan, J., Arnold, K. S., Taylor, S., Taylor, J. M., & Innerarity, T. L. (1995). Apolipoprotein B mRNA-editing protein induces

hepatocellular carcinoma and dysplasia in transgenic animals. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8483–8487. https://doi.org/10.1073/PNAS.92.18.8483

Yamanaka, S., Poksay, K. S., Arnold, K. S., & Innerarity, T. L. (1997). A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. *Genes & Development*, *11*(3), 321–333. https://doi.org/10.1101/GAD.11.3.321

Yuen, C. K., Lam, J. Y., Wong, W. M., Mak, L. F., Wang, X., Chu, H., Cai, J. P., Jin, D. Y., To, K. K. W., Chan, J. F. W., Yuen, K. Y., & Kok, K. H. (2020). SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. *Emerging Microbes and Infections*, *9*(1), 1418–1428. https://doi.org/10.1080/22221751.2020.1780953/SUPPL_FILE/TEMI_A_1780953_SM1758.ZIP

Zhang, H., Yang, B., Pomerantz, R. J., Zhang, C., Arunachalam, S. C., & Gao, L. (2003). The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature*, *424*(6944), 94–98. https://doi.org/10.1038/NATURE01707

Supplemental data

Primers for lentiviral plasmids construction

Table S1: Primers used for the construction of lentiviral vectors. These lentiviral vectors have been transduced in Vero E6 cells for SARS-CoV-2 replication kinetics assay.

Fragment	Insert source	Primer couple
A1 3' fragment	pcDNA3.1D_A1_V5_His_TOPO	catcgtatgggtagccgccctcgagtctccaagccacagaagg ccaaactcacctggtgtgactctagtgatc
A1 intron fragment	Plenti4_Blasti_A3B_HA	gtcacaccaggtgagtttggggacccttg agtcggcaccctgtaggaaagagaagaaggc
A1 5' fragment	pcDNA3.1D_A1_V5_His_TOPO	tttcctacagggtgccgactcagaaactc tgatagagaccgcgggccctctagaatgacttctgagaaaggtcc
A2 3' fragment	pcDNA3.1D_A2_V5_His_TOPO	catcgtatgggtagccgccctcgagcttcaggatgtctgccaacttc ccaaactcaccctgtgcagcgtgtgctg
A2 intron fragment	Plenti4_Blasti_A3B_HA	gctgcacagggtgagtttggggacccttg tcctccagccctgtaggaaagagaagaaggc
A2 5' fragment	pcDNA3.1D_A2_V5_His_TOPO	tttcctacagggctggaggacacatacc tgatagagaccgcgggccctctagaatggcccagaaggaagag
A3A 3' fragment	pcDNA3.1D_A3Ai4_V5-His-TOPO	catcgtatgggtagccgccctcgaggtttccctgattctggagaatggcccgcag ccaaactcaccctgcttctcctggggct
A3A intron fragment	Plenti4_Blasti_A3B_HA	gagaagcagggtgagtttggggacccttg tcctggagccctgtaggaaagagaagaaggc
A3A 5' fragment	pcDNA3.1D_A3Ai4_V5-His-TOPO	tttcctacagggctccaggagatgaaccaag tgatagagaccgcgggccctctagaatggaagccagcccagca
A3C 3' fragment	pcDNA3.1D_A3C_V5_His_TOPO	catcgtatgggtagccgccctcgagctggagactctcccgtag ccaaactcacctaacacaaagtaccaggtc
A3C intron fragment	Plenti4_Blasti_A3B_HA	tttgtgttaggtgagtttggggacccttg atactgtctcctgtaggaaagagaagaaggc
A3C 5' fragment	pcDNA3.1D_A3C_V5_His_TOPO	tttcctacaggagacagtatgtcgtcgc tgatagagaccgcgggccctctagaatgaatccacagatcagaaac
A3DE 3' fragment	pcDNA3.1D_A3DE_V5_His_TOPO	catcgtatgggtagccgccctcgagctggagaatctcccgtag ccaaactcacctggtacacatcttggag
A3DE intron fragment	PLENTI4_A3B_HA_Blasti	tgtgtaccaggtgagtttggggacccttg acgaggtcacctgtaggaaagagaagaaggc
A3DE 5' fragment	pcDNA3.1D_A3DE_V5_His_TOPO	tttcctacaggtgacctcgtagtttgtg tgatagagaccgcgggccctctagaatgaatccacagatcagaaatc
A3F 3' fragment	pcDNA3.1D_A3F_V5_His_TOPO	catcgtatgggtagccgccctcgagctcgagaatctcctgcag ccaaactcaccgatggaggcaatgtatc
A3F intron fragment	Plenti4_Blasti_A3B_HA	gcctccatcggtgagtttggggacccttg ctcagaaaccctgtaggaaagagaagaaggc
A3F 5' fragment	pcDNA3.1D_A3F_V5_His_TOPO	tttcctacagggtttctgagaatctcctttag tgatagagaccgcgggccctctagaatgaagcctcacttcagaaac
A3G 3' fragment	pcDNA3.1D_A3G_V5_His_TOPO	catcgtatgggtagccgccctcgaggttttcctgattctggag ccaaactcaccacattcactttcaactttaac
A3G intron fragment	Plenti4_Blasti_A3B_HA	agtgaatgtggtgagtttggggacccttg tggatccaccctgtaggaaagagaagaaggc
A3G 5' fragment	pcDNA3.1D_A3G_V5_His_TOPO	tttcctacagggtggatccatcgagtgtc tgatagagaccgcgggccctctagaatgaagcctcacttcagaaac

A3H 3' fragment	pcDNA3.1D_A3H _V5_His_TOPO	catcgtatgggtagccgccctcgagggactgctttatcctgtcaagc ccaaactcaccctgctcctcctgtgcct
A3H intron fragment	Plenti4_Blasti_A3 B_HA	gaggagcagggtgagtttggggacccttg acgtggagccctgtaggaaagagaagaaggc
A3H 5' fragment	pcDNA3.1D_A3H _V5_His_TOPO	tttcctacagggctccacgtgaggtaac tgatagagaccgcgggccctctagaatggctctgttaacagcc

Table S2: Primers used for the redesign and reconstruction of lentiviral vectors for the replication kinetics assay.

Fragment	**Insert source**	**Primer couple**
A3DE 3' fragment	pcDNA3.1D_A3D E_V5_His_TOPO	catcgtatgggtagccgccctcgagctggagaatctcccgtag ccaaactcacctggtacacatcttggag
A3DE intron fragment	PLENTI4_A3B_H A_Blasti	tgtgtaccaggtgagtttggggacccttg acgaggtcacctgtaggaaagagaagaaggc
A3DE 5' fragment	pcDNA3.1D_A3D E_V5_His_TOPO	tttcctacaggtgacctcgtagtttgtg tgatagagaccgcgggccctctagaatgaatccacagatcagaaatc
A3F 3' fragment	pcDNA3.1D_A3F_ V5_His_TOPO	catcgtatgggtagccgccctcgagctcgagaatctcctgcag ccaaactcacctggtacacatcttggag
A3F intron fragment	Plenti4_Blasti_A3 B_HA	tgtgtaccaggtgagtttggggacccttg acgaggtcacctgtaggaaagagaagaaggc
A3F 5' fragment	pcDNA3.1D_A3F_ V5_His_TOPO	tttcctacaggtgacctcgtagtttgtg tgatagagaccgcgggccctctagaatgaagcctcacttcagaaac
A3G 3' fragment	pcDNA3.1D_A3G _V5_His_TOPO	catcgtatgggtagccgccctcgaggttttcctgattctggag ccaaactcacctgtgttatgaggtggag
A3G intron fragment	Plenti4_Blasti_A3 B_HA	cataacacaggtgagtttggggacccttg tgagacttacctgtaggaaagagaagaaggc
A3G 5' fragment	pcDNA3.1D_A3G _V5_His_TOPO	tttcctacaggtaagtctcatgccgtcc tgatagagaccgcgggccctctagaatgaagcctcacttcagaaac

Table S3: Primers used for the construction of the ShA3G lentiviral vector.

Fragment	**Insert source**	**Primer couple**
shRNA A3G	/	tgcatcgtgaccaggagtatttcaagagaatactcctggtcacgatgctttttc tcgagaaaaaagcatcgtgaccaggagtattctcttgaaatactcctggtcacgatgca
Blasti fragment	pSicoR_EF1a_mC herry_T2A_Blasti_ sh_scrambled	attacccgggagggcagaggaagtc atgagaattcttagccctcccacacataacc

Maps of lentiviral plasmids

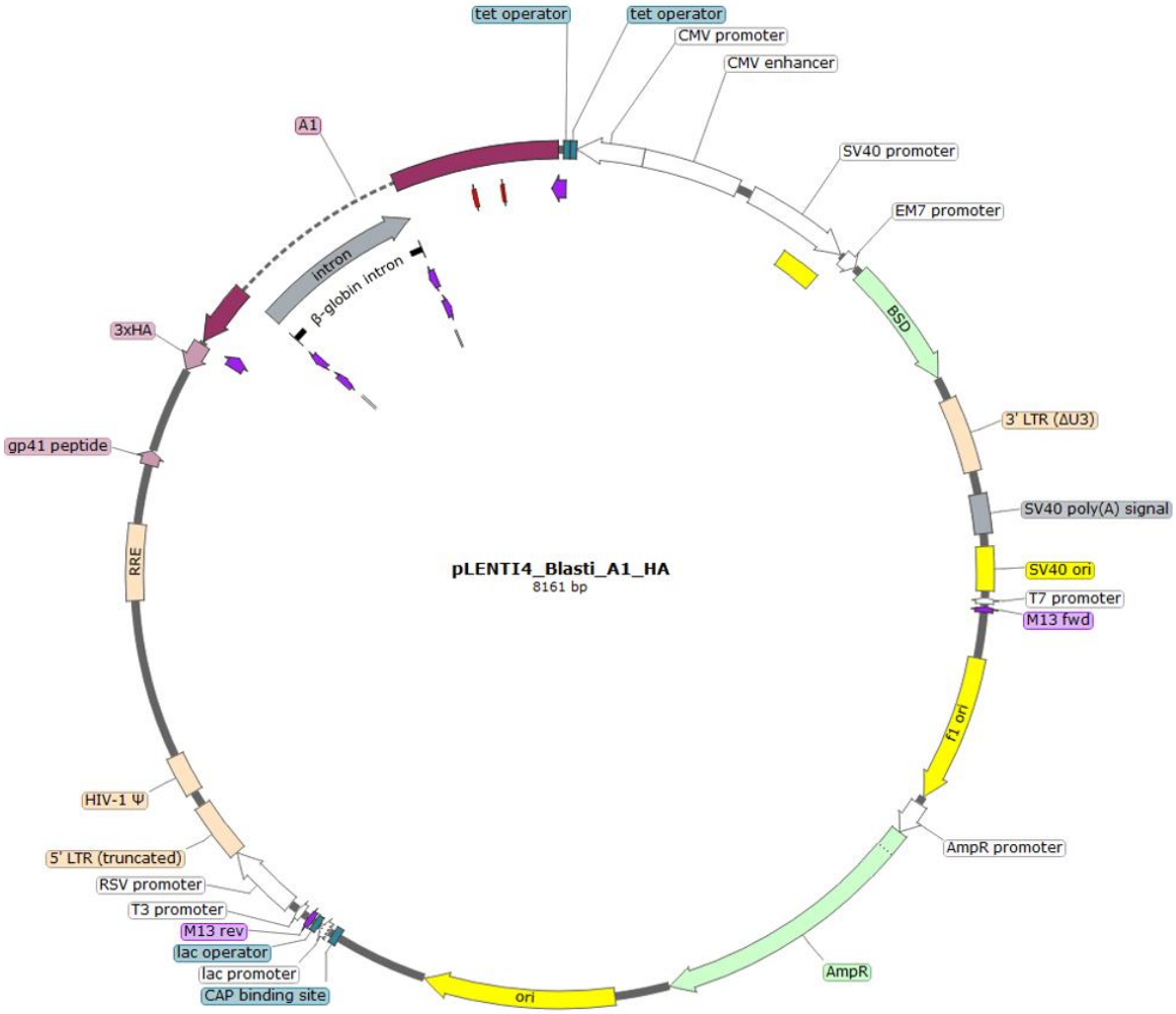

Figure S1: Map of the lentiviral plasmid bearing Ha-tagged A1 transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene.

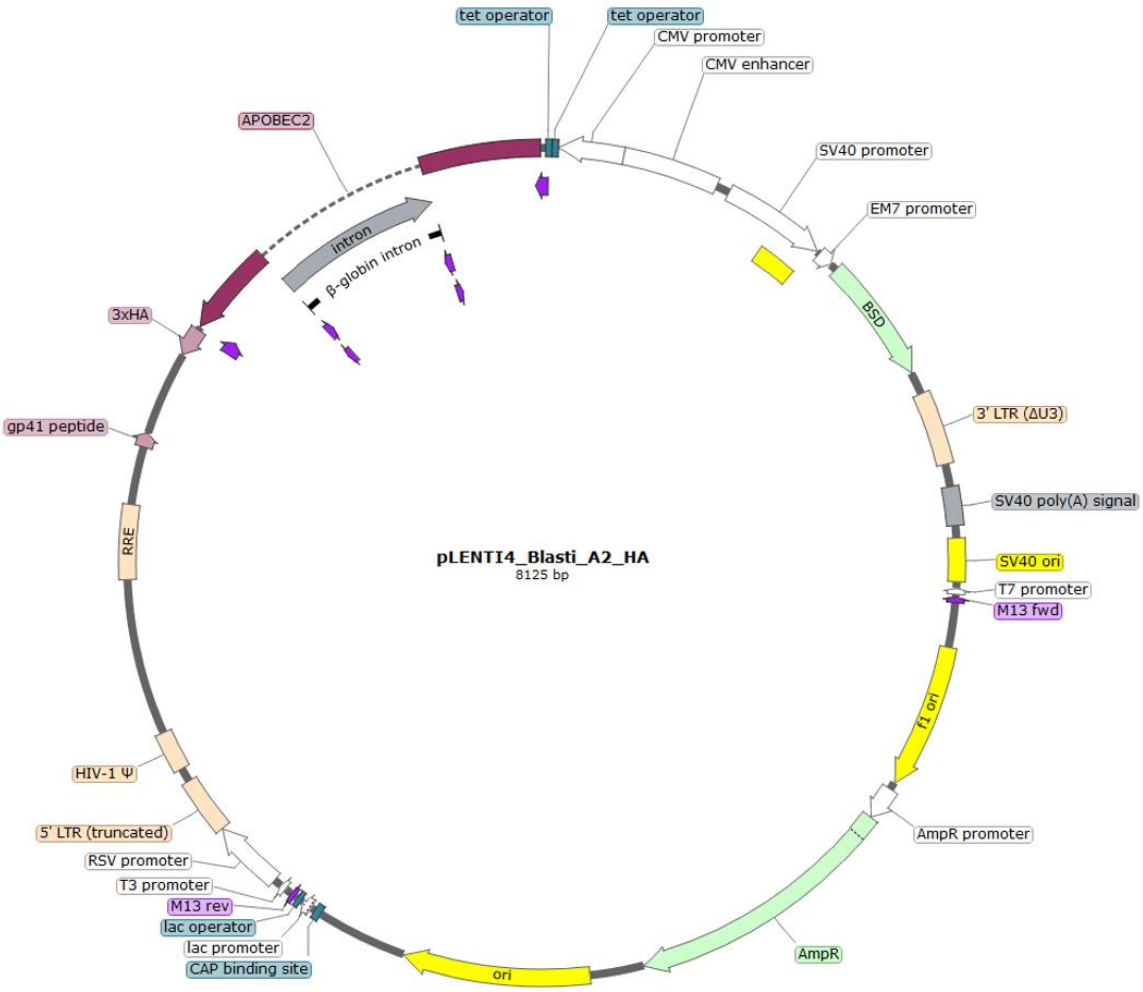

Figure S2: Map of the lentiviral plasmid bearing Ha-tagged A2 transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene.

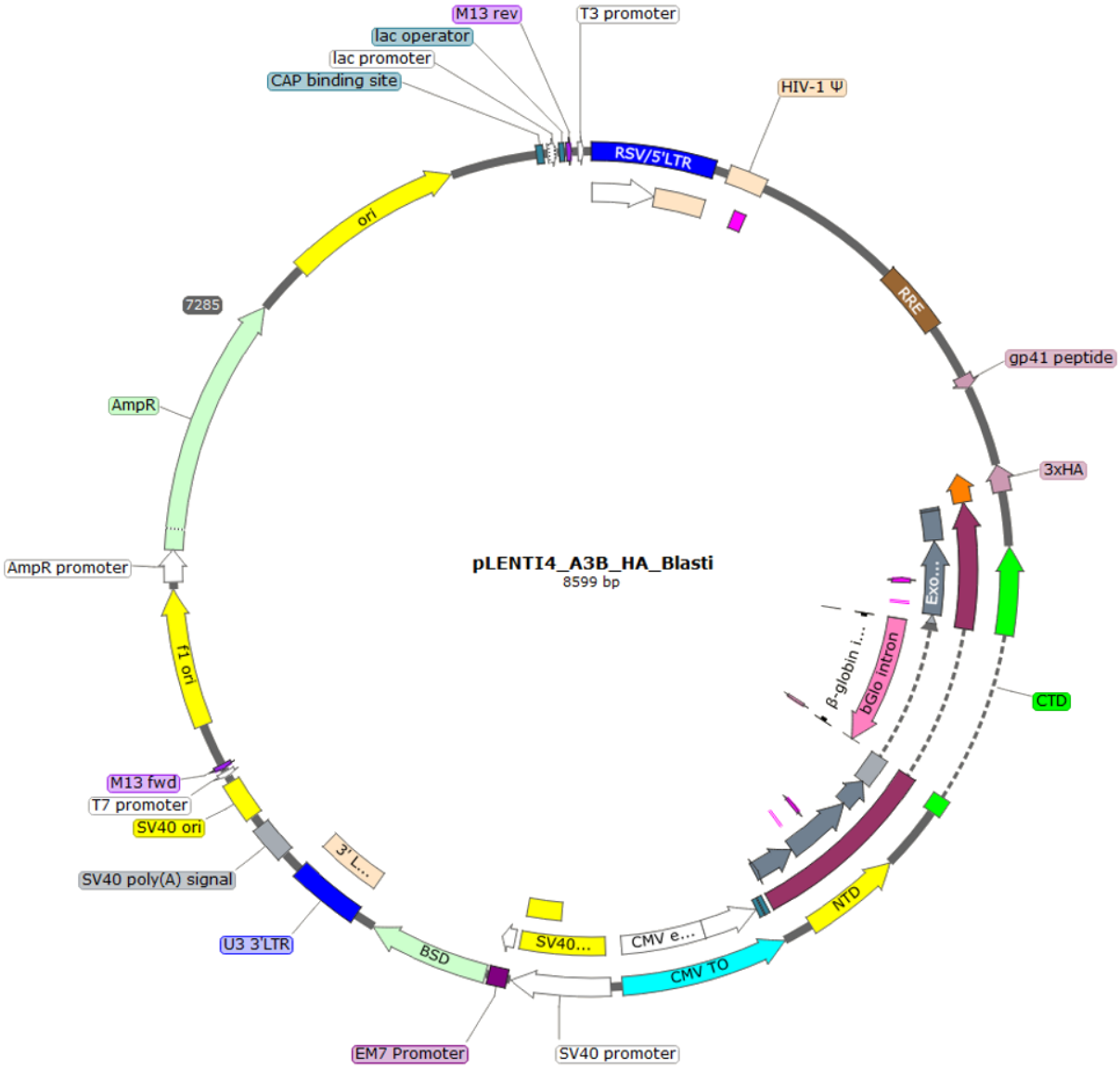

Figure S3: Map of the lentiviral plasmid bearing Ha-tagged A3B transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene.

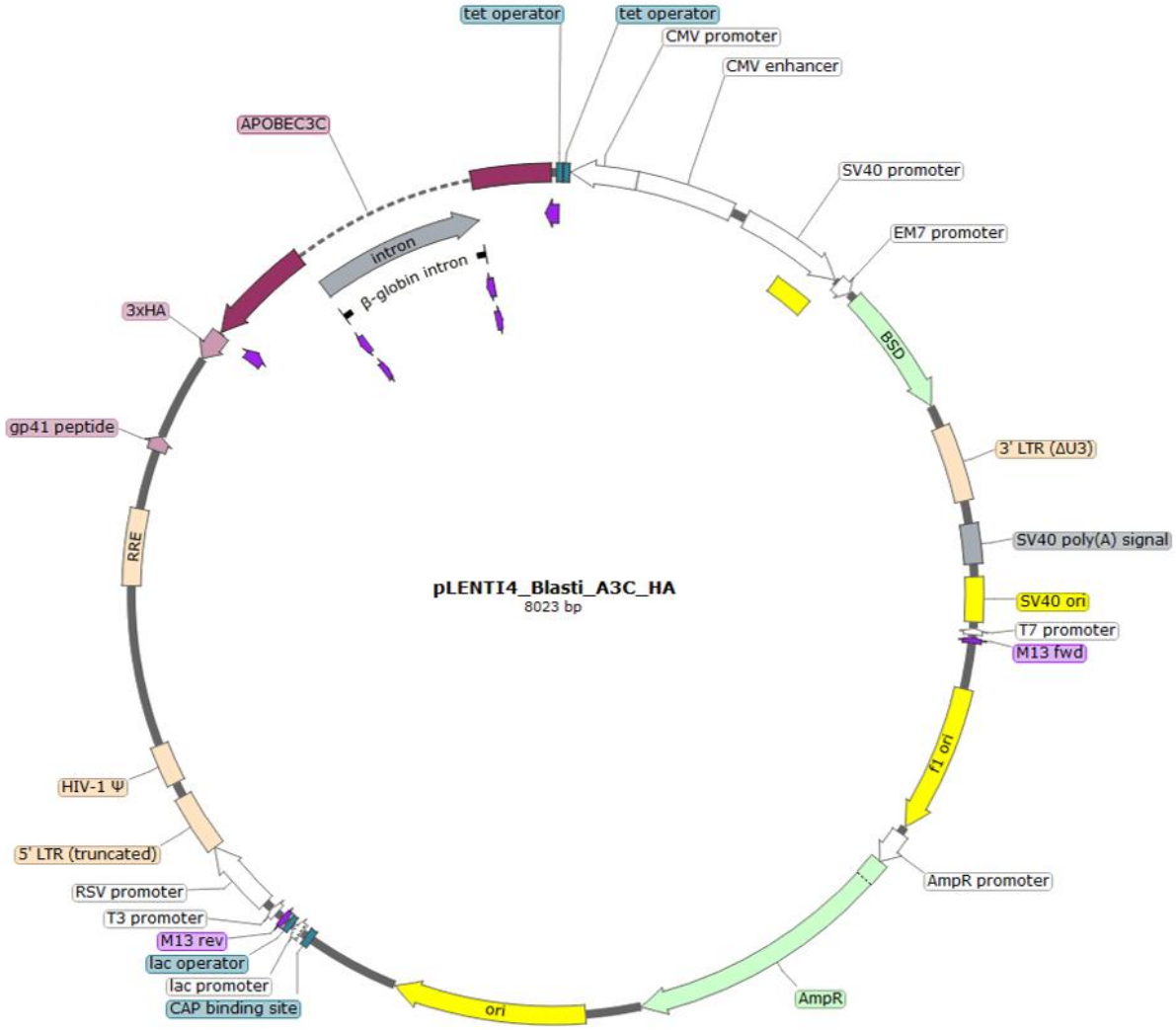

Figure S4: Map of the lentiviral plasmid bearing Ha-tagged A3C transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene.

Figure S5: Map of the lentiviral plasmid bearing Ha-tagged A3DE transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene.

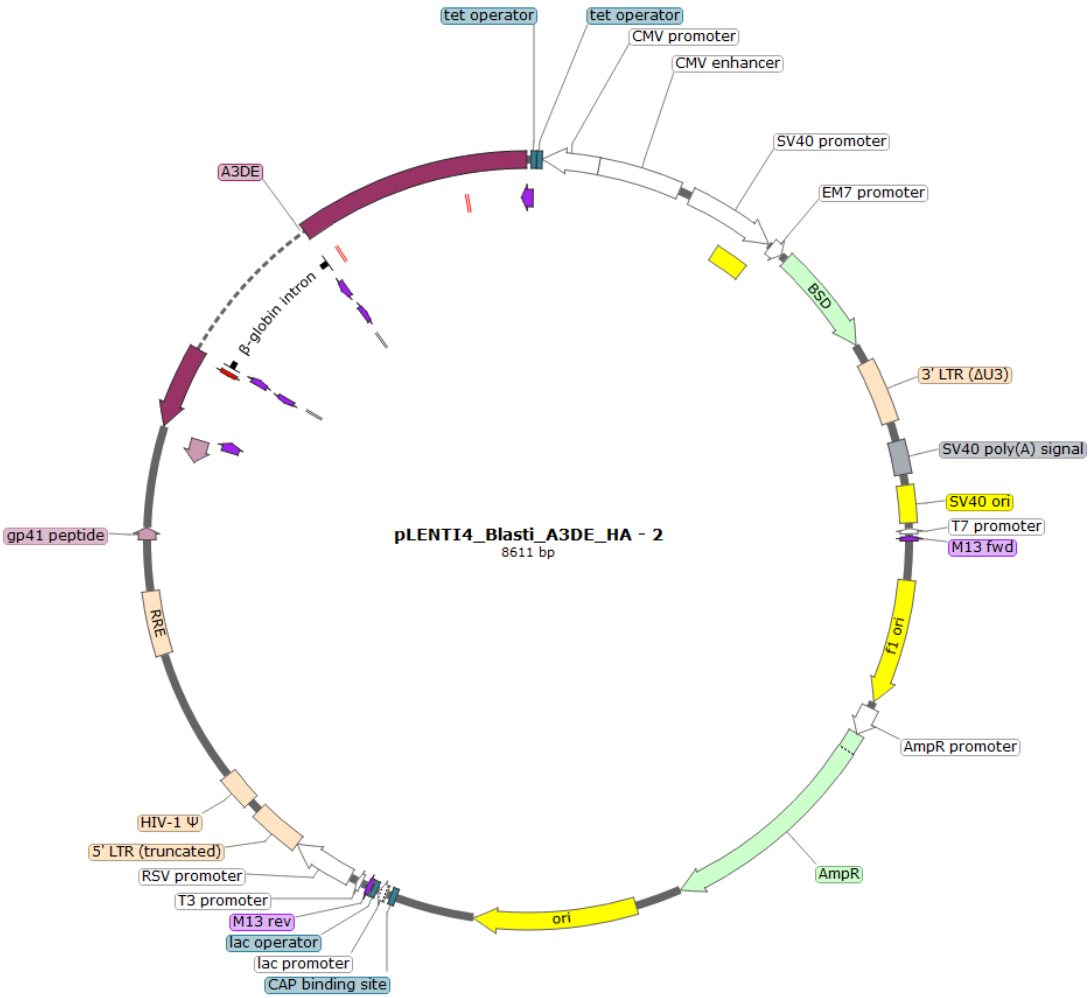

Figure S6: Map of the lentiviral plasmid bearing the second version of Ha-tagged A3DE transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene. This plasmid is the redesigned version of the A3DE lentiviral vector that will be used in further assays.

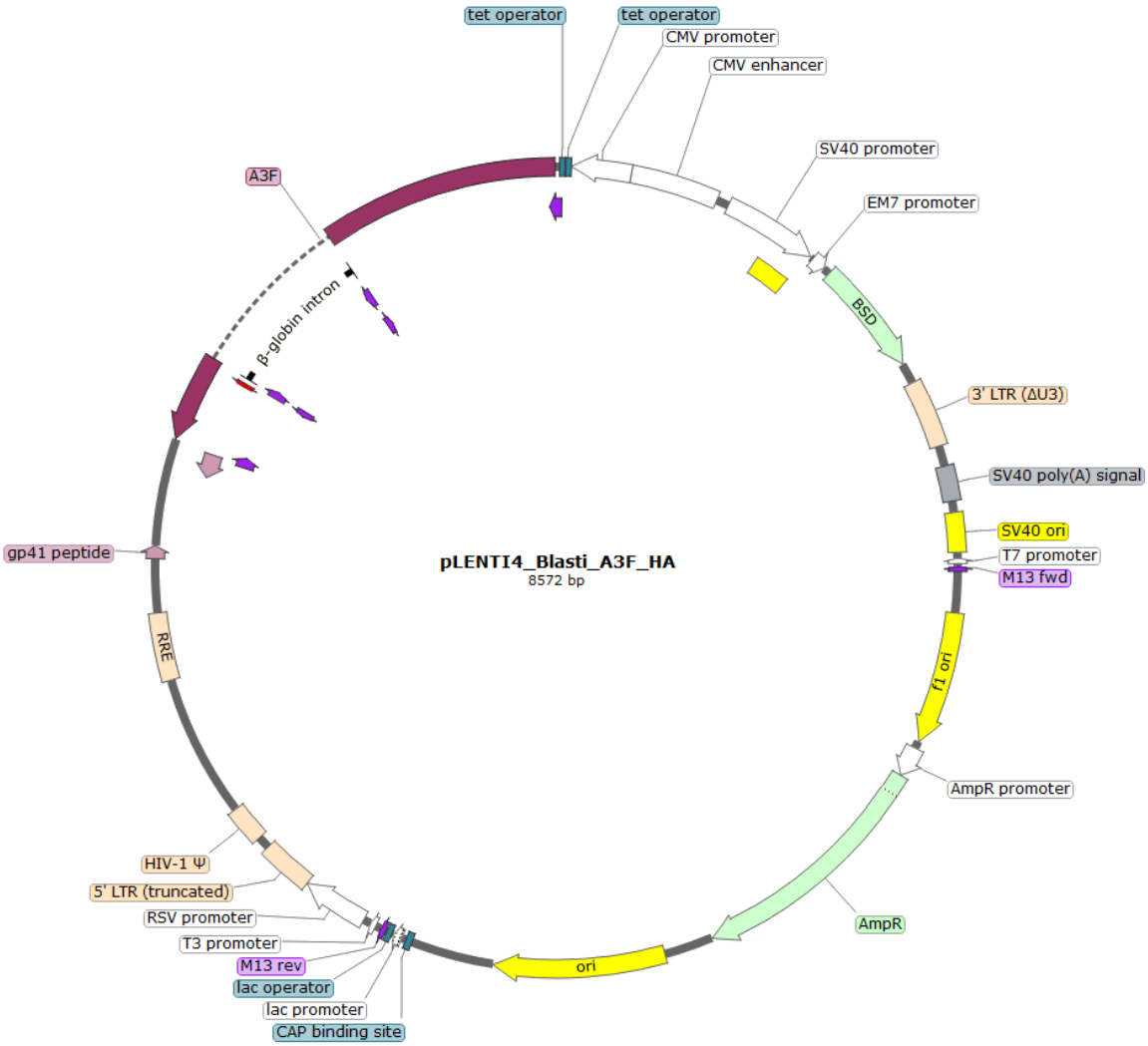

Figure S7: Map of the lentiviral plasmid bearing Ha-tagged A3F transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene.

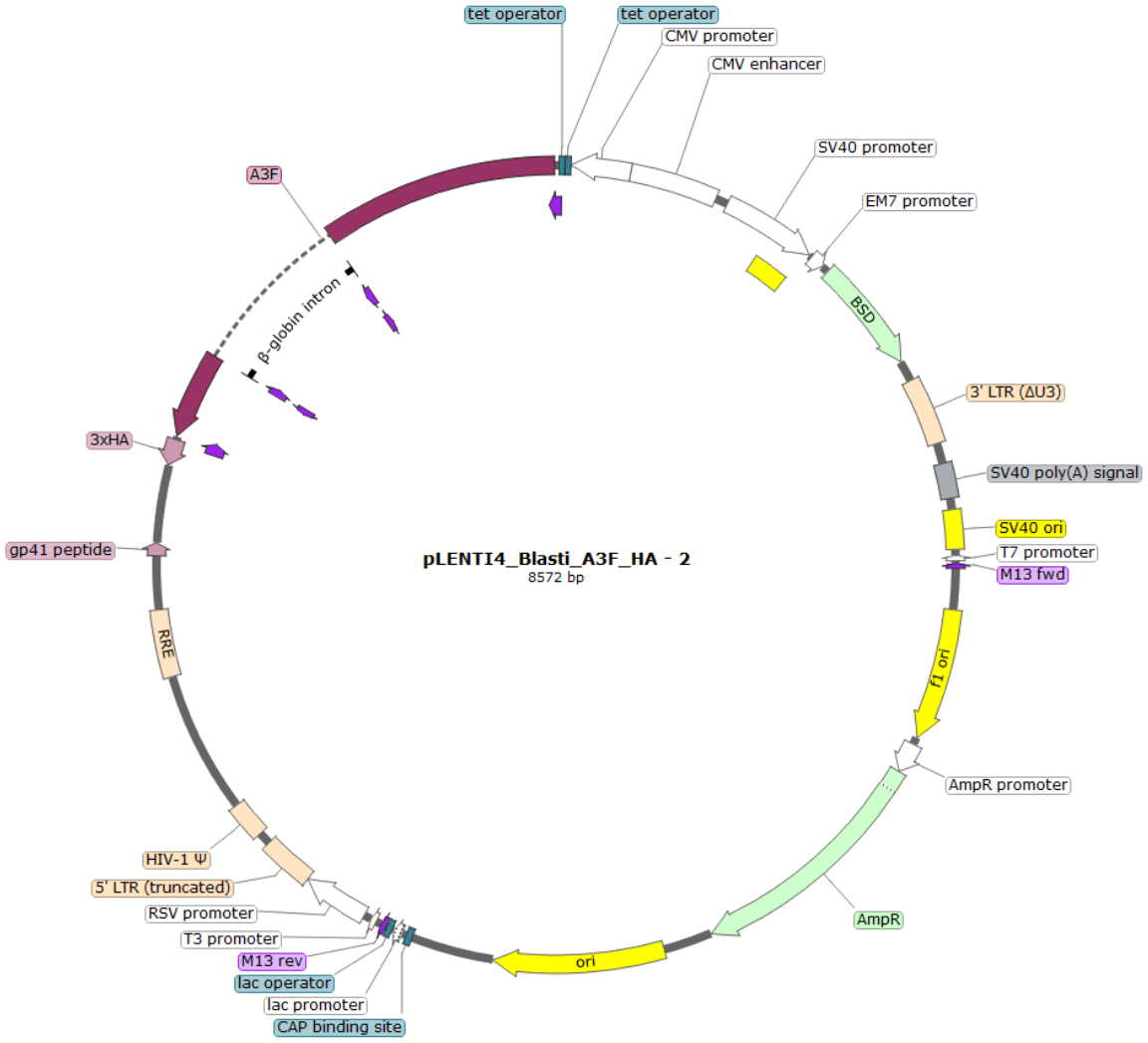

Figure S8: Map of the lentiviral plasmid bearing the second version of Ha-tagged A3F transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene. This plasmid is the redesigned version of the A3F lentiviral vector that will be used in further assays.

Figure S9: Map of the lentiviral plasmid bearing Ha-tagged A3G transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene.

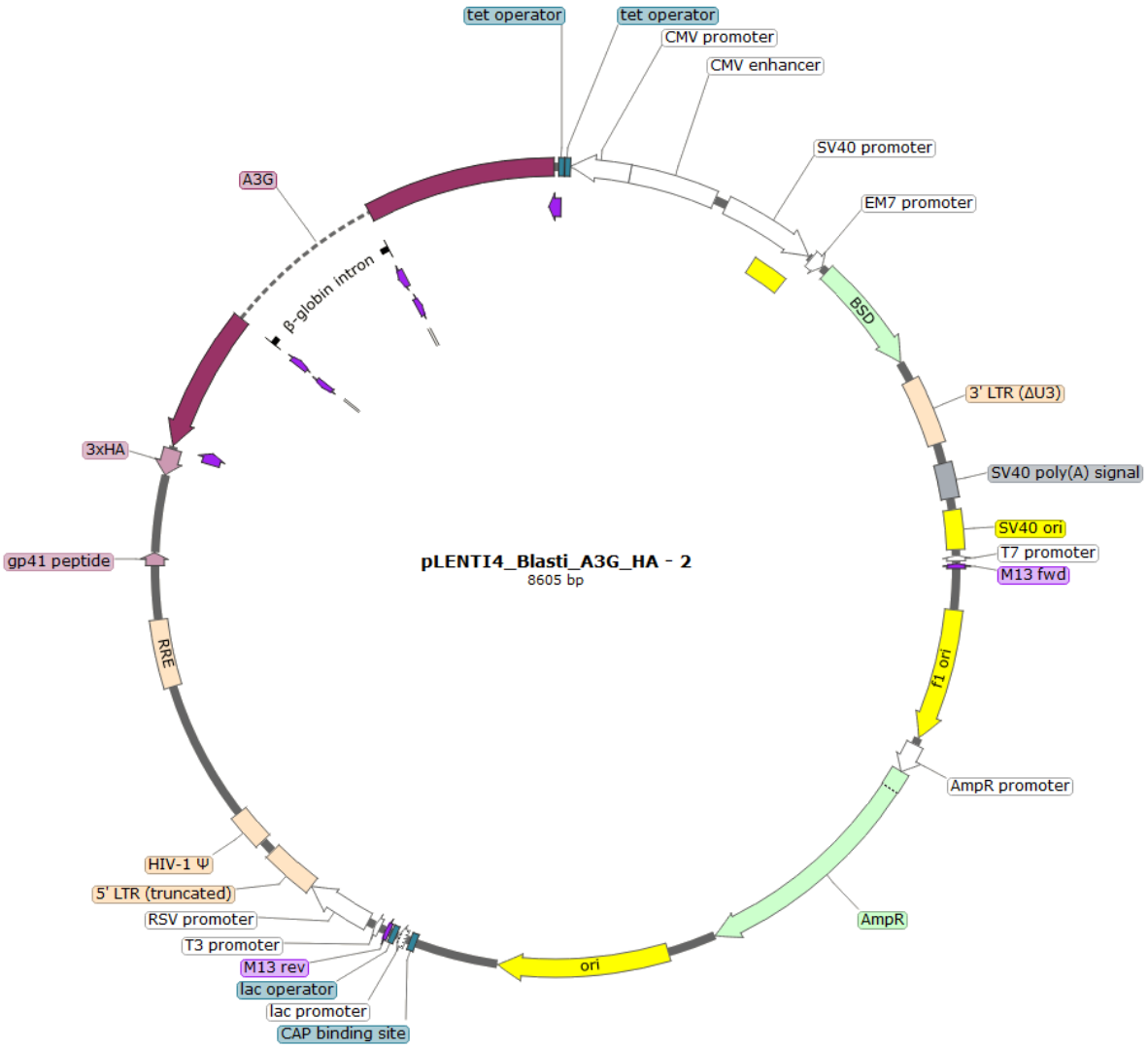

Figure S10: Map of the lentiviral plasmid bearing the second version of Ha-tagged A3G transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene. This plasmid is the redesigned version of the A3G lentiviral vector that will be used in further assays.

Figure S11: Map of the lentiviral plasmid bearing Ha-tagged A3H transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production. Purple arrows show the hybridization site of primers for the cloning of the transgene.

Figure S12: Map of the lentiviral plasmid bearing GFP transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production.

Figure S13: Map of the lentiviral plasmid bearing hACE2 transgene and hygromycin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production.

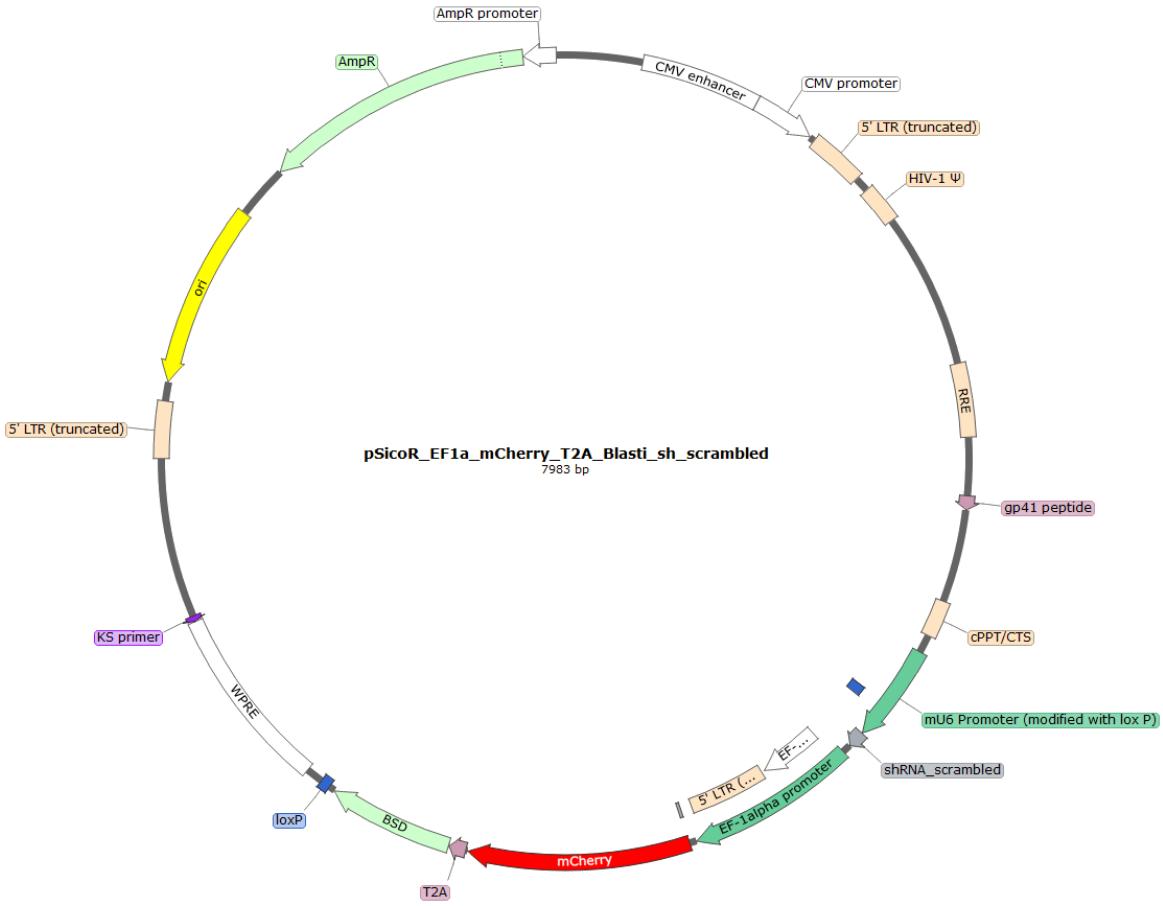

Figure S14: Map of the lentiviral plasmid bearing the scrambled shRNA transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production.

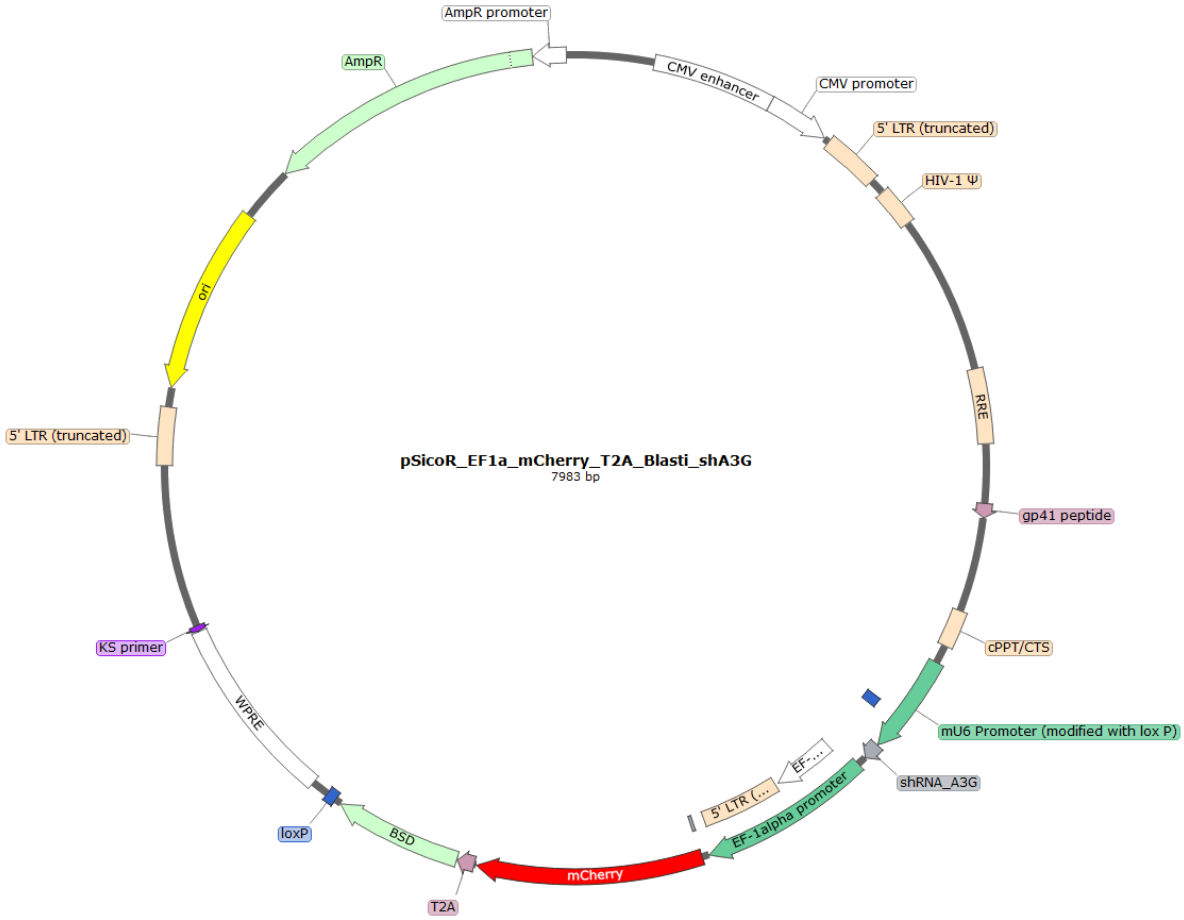

Figure S15: Map of the lentiviral plasmid bearing the scrambled shRNA A3G transgene and blasticidin resistance gene. The plasmid is used for transfection of HEK 293T cells and lentivirus production.

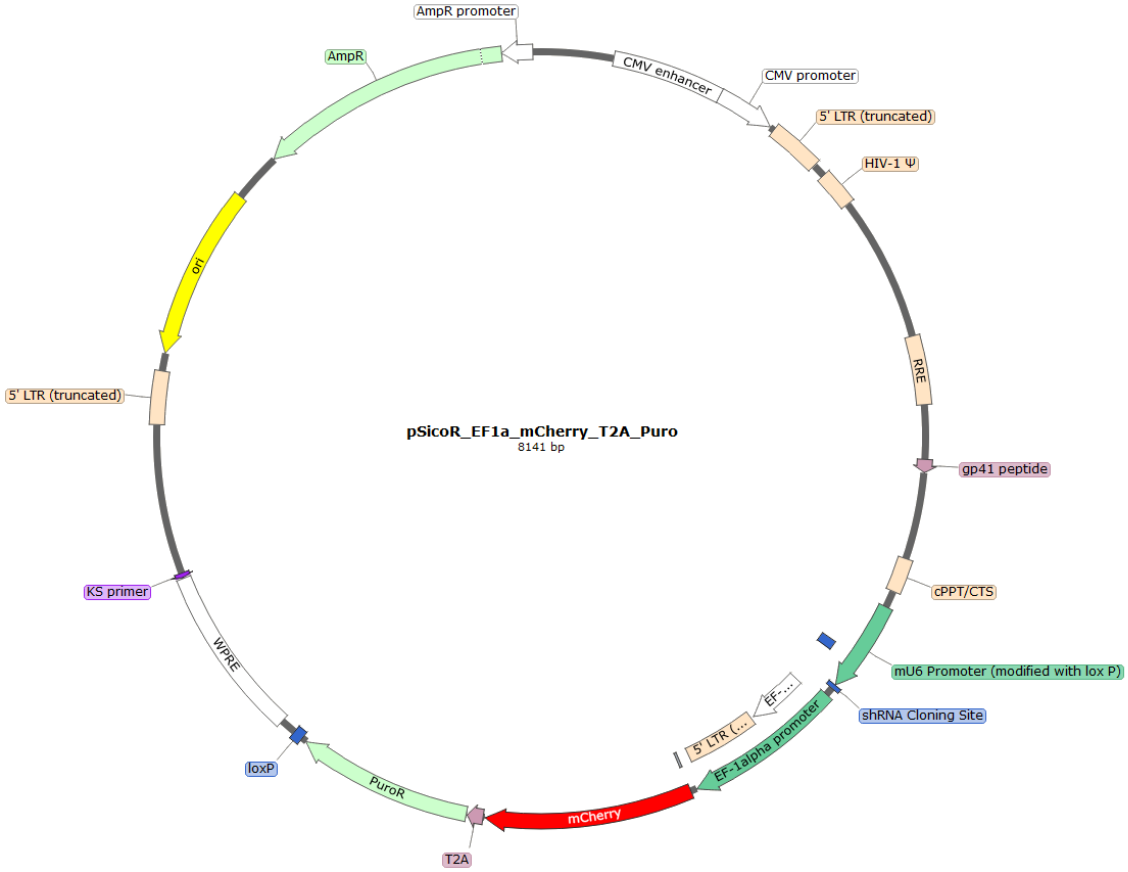

Figure S16: Map of the lentiviral plasmid bearing the shRNA cloning site and puromycin resistance gene. The plasmid is used for construction of ShA3G plasmid.

Thermal cycler protocol for RTqPCR

Table S4: Thermal cycler protocol and primers used for viral RNA quantification by RTqPCR. Fam: 6-carboxyfluorescein. BHQ1: Black Hole Quencher®-1

Step	**Protocol**	
Reverse transcription	48°C –10min	
Initialization	95°C – 3min	
Denaturation	95°C – 15sec	45 cycles
Annealing / Extension	58°C – 30sec	
Primers		
E_Sarbeco_Fw	5’ - ACAGGTACGTTAATAGTTAATAGCGT-3’	
E_Sarbeco_Rev	5’ - ATATTGCAGCAGTACGCACACA-3’	
E_Sarbeco_Probe	5’ -(FAM) ACACTAGCCATCCTTACTGCGCTTCG(BHQ1)-3’	

Table S5: Thermal cycler protocol and primers used for RNA quantification of APOBEC mRNA by RTqPCR.

Step	**Protocol**	
Initialization	95°C – 3min	45 cycles
Denaturation	95°C – 10sec	
Annealing	58°C – 20sec	
Extension	72°C – 20sec	
Primers		
TBP_Fw	ACCTAAAGACCATTGCACTTCG	
TBP_Rv	CATATTTTCTTGCTGCCAGTCTG	
GAPDH_Fw	ATTCCCATCACCATCTTCCAG	
GAPDH_Rv	CAGAGATGATGACCCTTTTGG	
HPRT_Fw	GGTCAGGCAGTATAATCCAAAG	
HPRT_Rv	AAGGGCATATCCTACAACAAAC	
A3A_Fw	GAGAAGGGACAAGCACATGG	
A3A_Rv	TGGATCCATCAAGTGTCTGG	
A3B_Fw	GACCCTTTGGTCCTTCGAC	
A3B_Rv	GCACAGCCCCAGGAGAAG	
A3C_Fw	AGCGCTTCAGAAAAGAGTGG	
A3C_Rv	AAGTTTCGTTCCGATCGTTG	
A3DE_Fw	ACCCAAACGTCAGTCGAATC	
A3DE_Rv	CACATTTCTGCGTGGTTCTC	
A3F_Fw	CCGTTTGGACGCAAAGAT	
A3F_Rv	CCAGGTGATCTGGAAACACTT	
A3G_Fw	CCGAGGACCCGAAGGTTAC	
A3G_Rv	TCCAACAGTGCTGAAATTCG	
A3H_Fw	AGCTGTGGCCAGAAGCAC	
A3H_Rv	CGGAATGTTTCGGCTGTT	

SARS-CoV2 genome

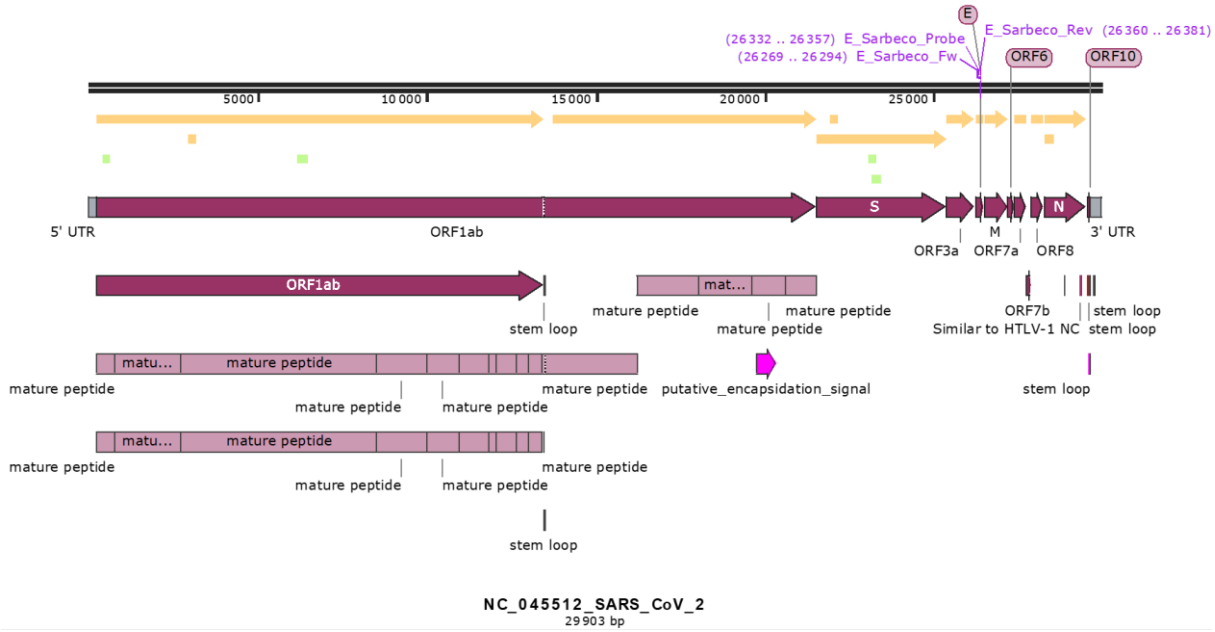

Figure S17: Schematic representation of the genome of the SARS-CoV-2 strain used for Vero E6 and HBEC3-KT-ACE2 infections. The strain was isolated from a Belgian patient infected with SARS-CoV-2 in March 2020.

Quantification of APOBEC proteins mRNA levels relative to housekeeping genes

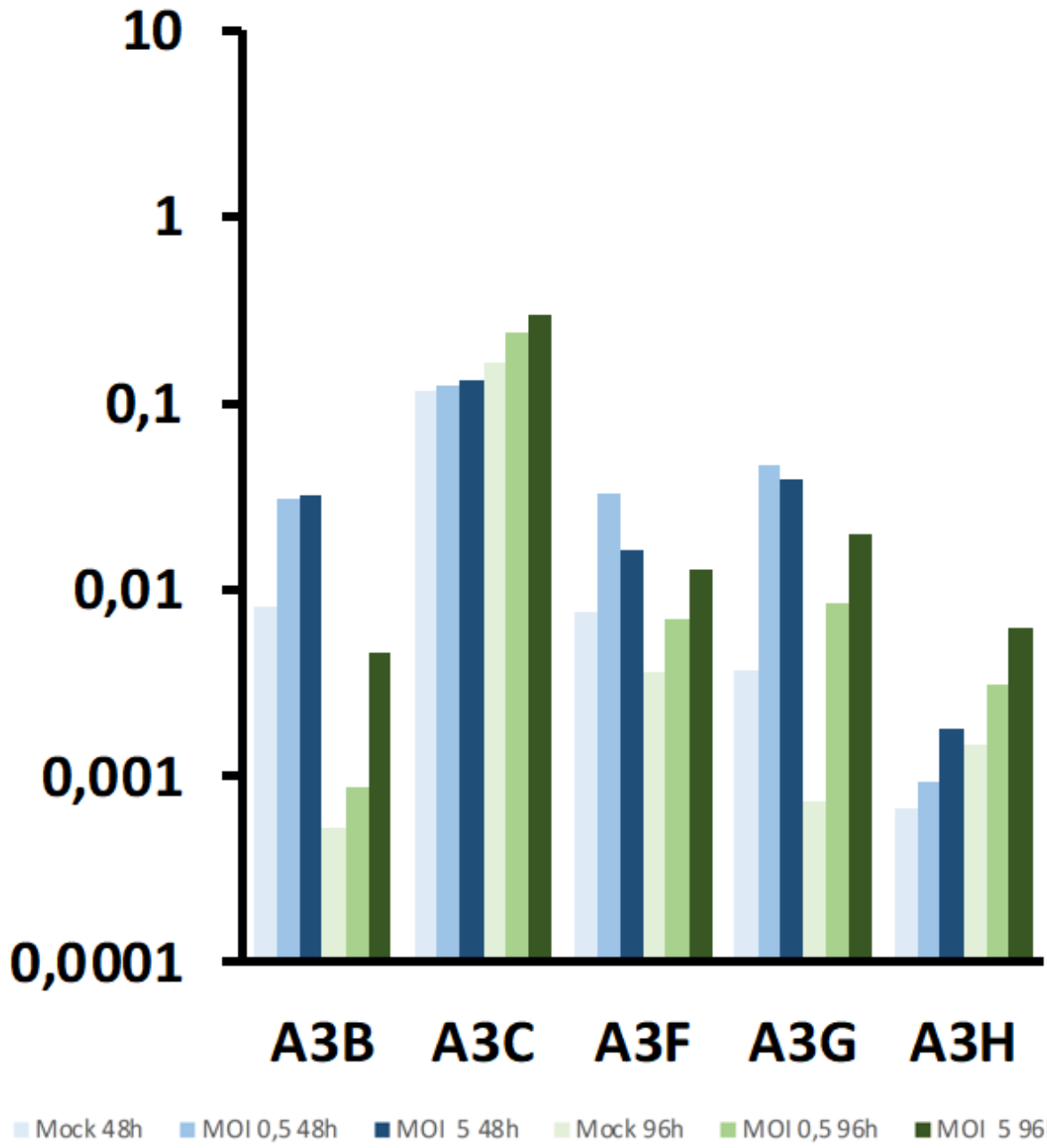

Figure S18: RNA quantification of A3 gene by RTqPCR upon HBEC3-KT-ACE2 infection by SARS-CoV-2 relative to housekeeping genes (HPRT/GAPDH/TBP) mRNA levels in replicate n°1. A3A and A3DE mRNAs were not detectable.

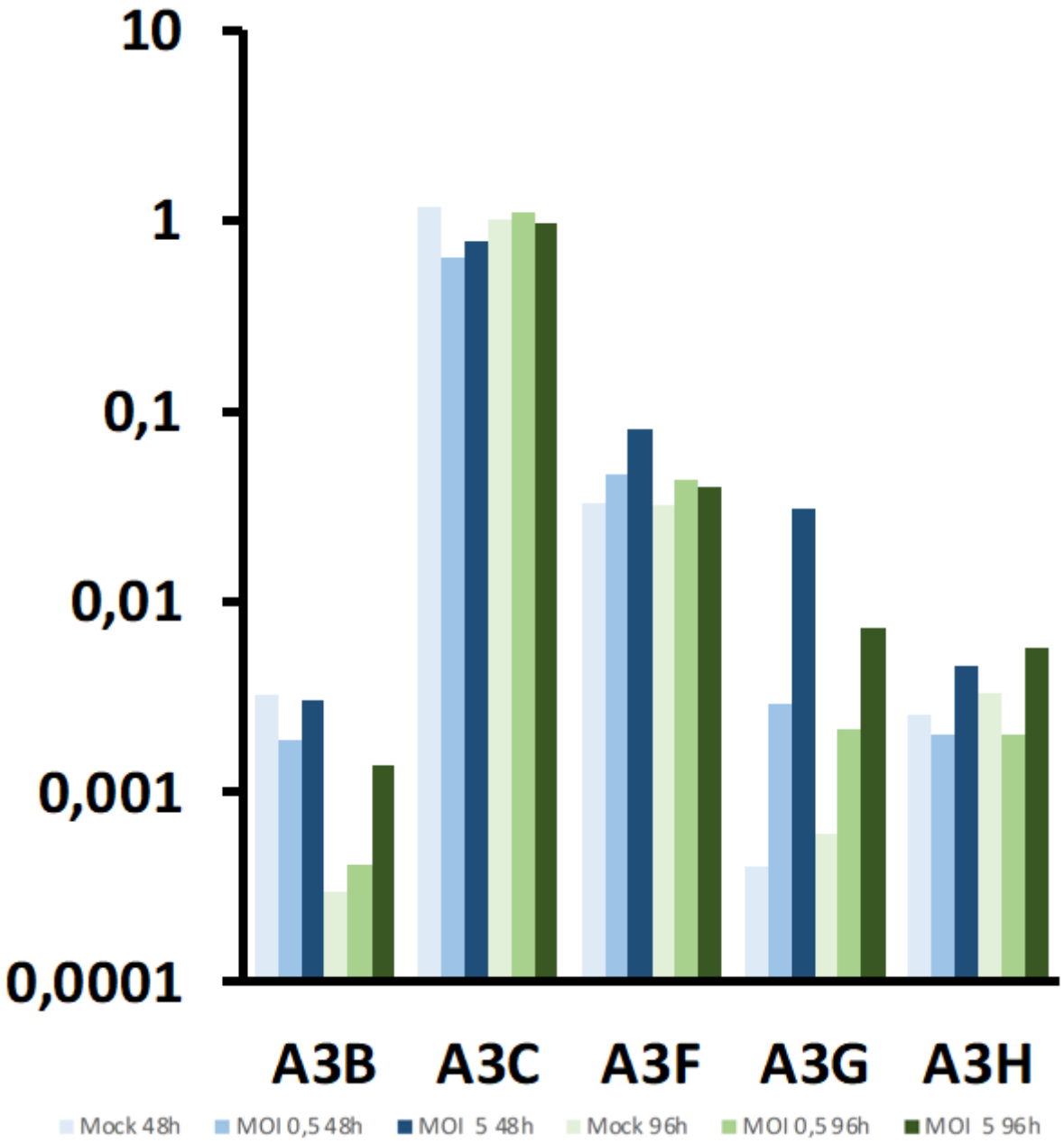

Figure S19: RNA quantification of A3 gene by RTqPCR upon HBEC3-KT-ACE2 infection by SARS-CoV-2 relative to housekeeping genes (HPRT/GAPDH/TBP) mRNA levels in replicate n°2. A3A and A3DE mRNAs were not detectable.

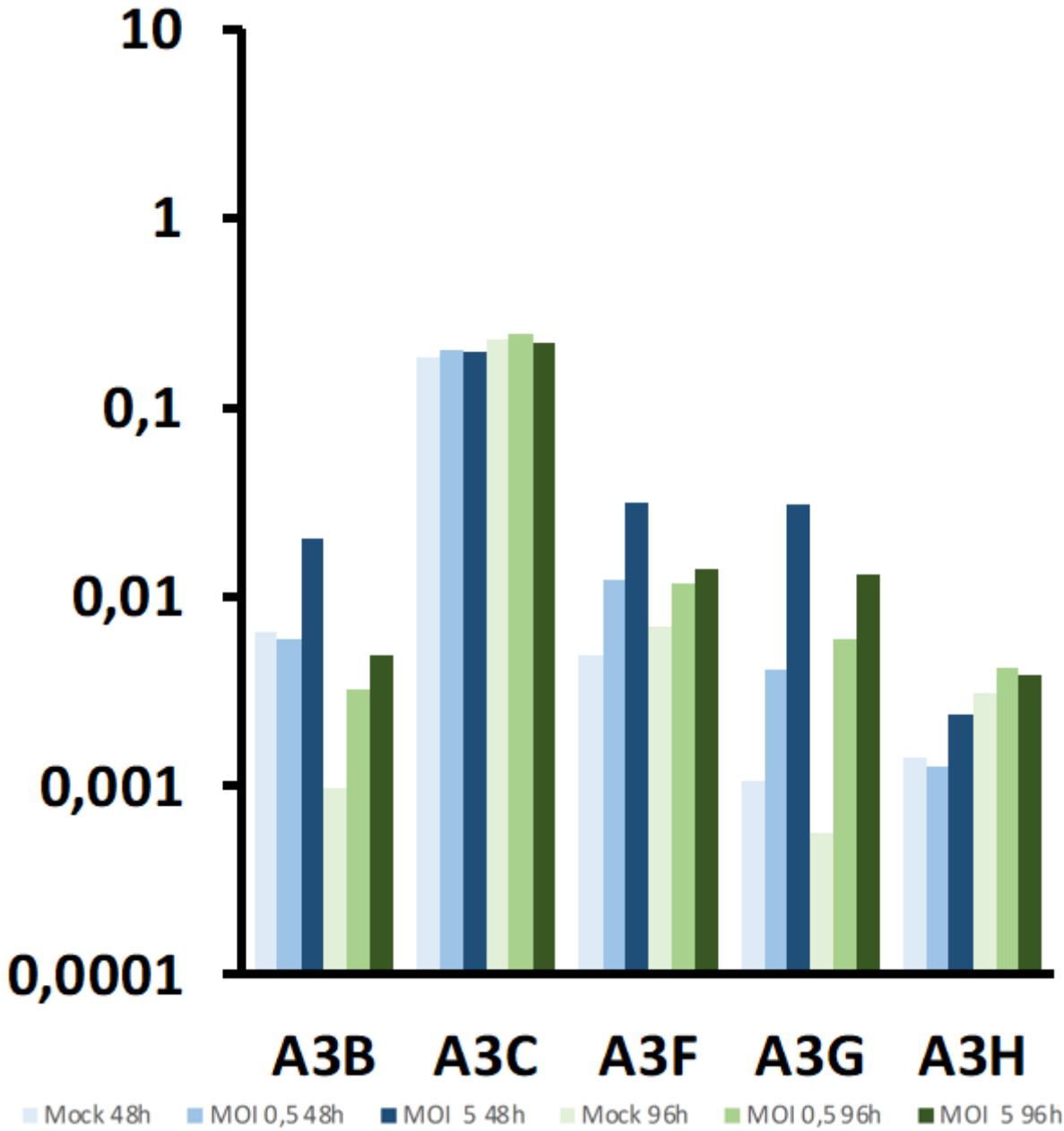

Figure S20: RNA quantification of A3 gene by RTqPCR upon HBEC3-KT-ACE2 infection by SARS-CoV-2 relative to housekeeping genes (HPRT/GAPDH/TBP) mRNA levels in replicate n°3. A3A and A3DE mRNAs were not detectable.

Assessment of A3G by Western blot in two others replicates

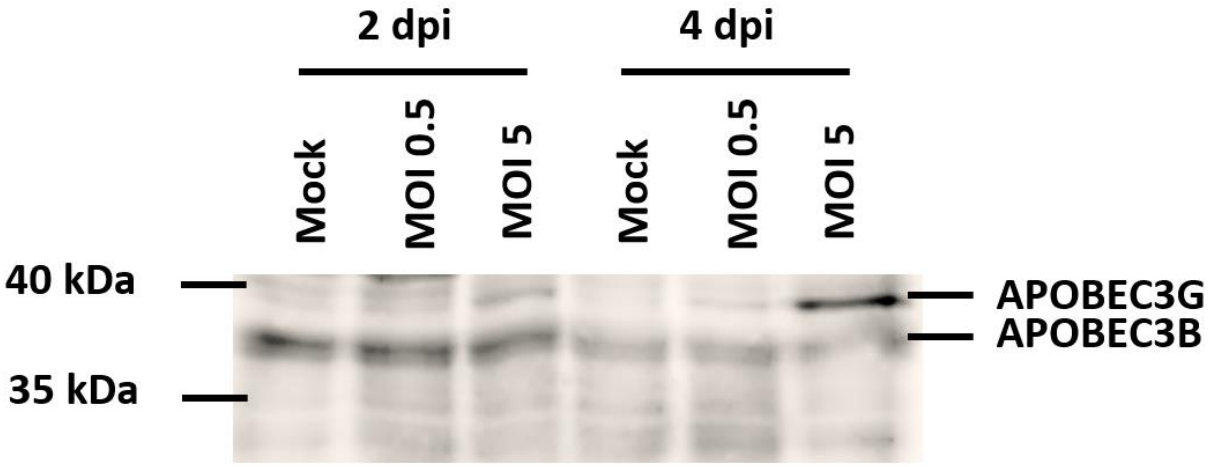

Figure S21: Assessment of A3G upon HBEC3-KT-ACE2 infection by SARS-CoV-2 in replicate n°2.

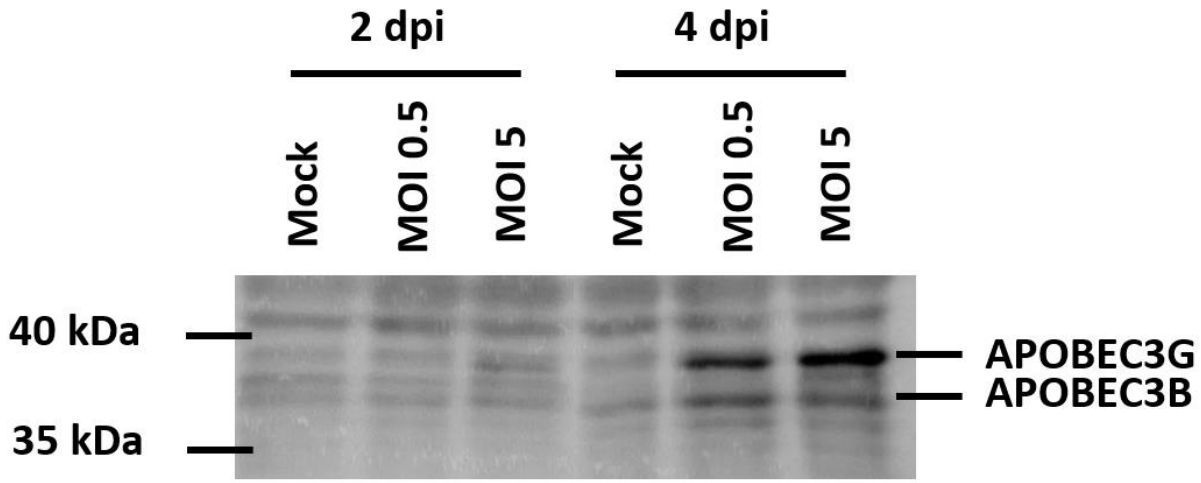

Figure S22: Assessment of A3G upon HBEC3-KT-ACE2 infection by SARS-CoV-2 in replicate n°3.

Assessment of A3G deaminase activity in two others replicates

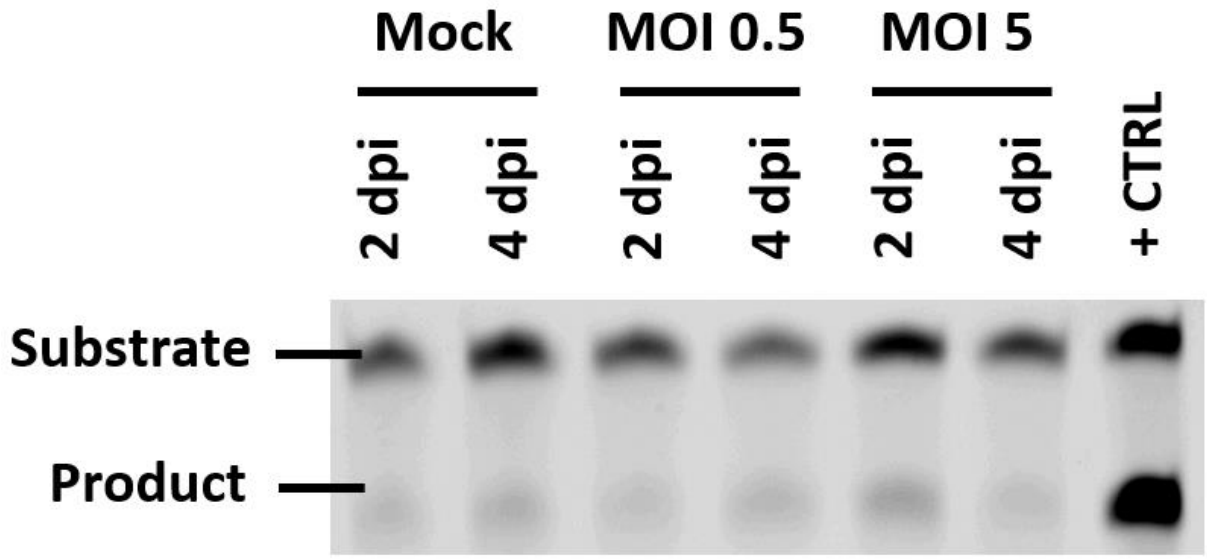

Figure S23: Assessment of A3G deaminase activity upon HBEC3-KT-ACE2 infection by SARS-CoV-2 in replicate n°1

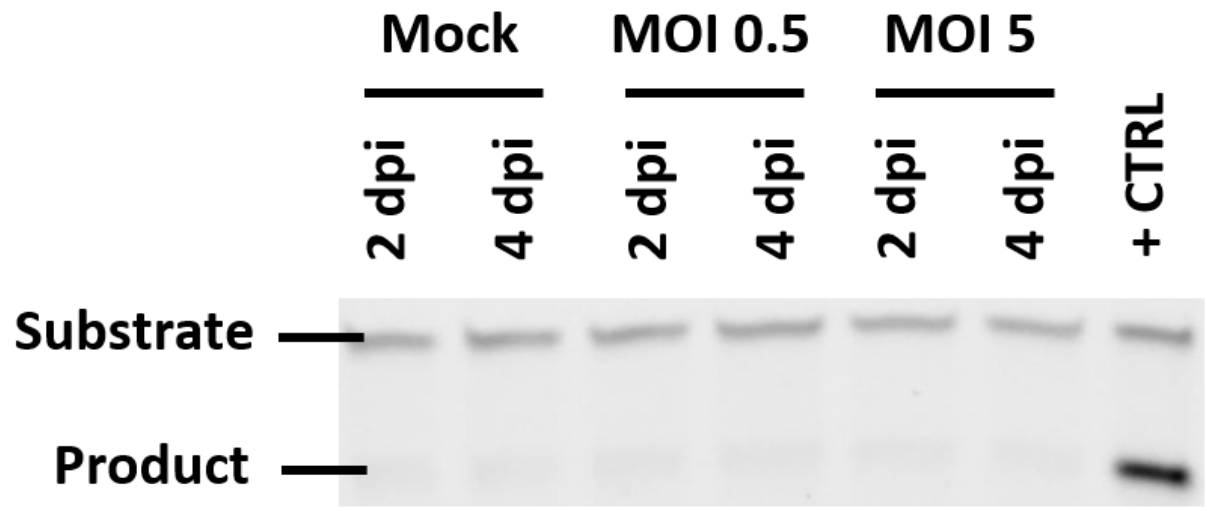

Figure S24: Assessment of A3G deaminase activity upon HBEC3-KT-ACE2 infection by SARS-CoV-2 in replicate n°3.