Charles University

Faculty of Science

Developmental and Cell Biology

MSc. Petr Nickl

Genome Engineering and Diseases Modelling Enhanced by Recombinant AAV Technology

Editace genomu *in vivo* pro modelování lidských onemocnění pomocí technologie AAV

Doctoral thesis

Supervisor:

PD. Dr. rer.nat. habil Radislav Sedlacek

Dept. of Transgenic Models of Diseases; Institute of Molecular Genetics of the Czech Academy of Sciences

Prague, 2024

I hereby declare that this thesis has been authored by me, Petr Nickl, and all literature sources have been properly cited. I confirm that the content of this dissertation has not been used for obtaining another academic degree or equivalent.

I hereby declare that artificial intelligence (AI) tools were used solely for the purpose of proofreading this thesis, including grammar correction and improving readability. No AI-generated content was used to develop the scientific ideas, concepts, or results presented in this work. All research, analysis, and intellectual contributions are my own.

Tímto prohlašuji, že tuto práci jsem napsal já, Petr Nickl, a veškerá použitá literatura byla řádně citována. Potvrzuji, že obsah této disertační práce nebyl využit k získání stejného nebo podobného titulu.

Tímto prohlašuji, že nástroje umělé inteligence (AI) byly použity výhradně pro účely korektury této práce, včetně opravy gramatiky a zlepšení čtivosti. Žádný obsah generovaný umělou inteligencí nebyl použit při vývoji vědeckých myšlenek, konceptů nebo výsledků prezentovaných v této práci. Veškerý výzkum, analýzy a intelektuální přínosy jsou mou vlastní prací.

Prague, October 2024

Acknowledgments

I would like to express my gratitude to my supervisor, *PD. Dr. rer.nat. habil Radislav Sedlacek*, for his invaluable guidance and support throughout the duration of this project. His expertise and encouragement have been instrumental in shaping my research and completing this dissertation.

A special thanks goes to my wife and family, whose unwavering support and understanding have been a source of strength and motivation during the most challenging moments of this journey.

I would also like to extend my heartfelt thanks to *Prof. Marjo Yliperttula*, who supervised my work at the University of Helsinki. Her mentorship and insights have greatly contributed to the progress of this research.

I am deeply grateful to my colleagues in the *Transgenic and Archiving module*, without whom this work would not have been possible. Their collaboration, dedication, and friendship have made this journey an incredibly rewarding experience.

List of abbreviations

AAP	Assembly-Activating Protein
AAV	adeno-associated virus (general term reffering to both, rAAV and wtAAV)
AAVR	AAV receptor
AAVS1	Adeno-associated virus integration site 1
AB	Apoptotic bodies
Ace2	mouse Angiotensin-converting enzyme 2
ACE2	human Angiotensin-converting enzyme 2
AdV	Adenovirus
ALIX	ALG-2-interacting Protein X
ATM	protein kinase ataxia telangiectasia mutated
ATP	adenosine triphosphate
BBB	blood brain barrier
BEVS	Baculovirus expression vector system
BR	basic region
Cap	AAV gene coding for capsid proteins
CD63	tetraspanin CD63
CD81	tetraspanin CD81
CD9	tetraspanin CD9
Cdc25A	cell division cycle 25 homolog A
CDS	gene coding sequences
CEP90	Centrosomal Protein Of 90 KDa
Chmp4	Charged Multivesicular Body Protein 4A
CMV	cytomegalovirus
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPR-READI	CRISPR RNP electroporation and AAV donor infection
dCas9	dead/inactive Cas9
dgRNA	dead gRNG
DMD	Duchenne muscular dystrophy
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
E1A	Adenovirus early region 1A
E2A	Adenovirus early region 2A
E2B	Adenovirus early region 1B
E2F-1	Transcription factor E2F1
E4orf6	Open reading frame 6 of early region 4
EASI-CRISPR	efficient additions with ssDNA inserts–CRISPR
EE	early endosome
EF1a	elongation factor 1 alpha
EGFP	enhanced green fluorescent protein
EMA	European Medicines Agency

ERAD	endoplasmic reticulum-associated protein degradation
ESCRT	endosomal sorting complexes required for transport
EUCOMM/KOMP	International Knockout Mouse Consortium
EV	extracellular vesicles
EV-AAV	extracellular vesicles associated with AAV
FDA	U.S. Food and Drug Administration
FI	forced inhalation
GC	genome copy/copies
GOI	gene of interest
GPCR	G protein-coupled receptor
GPR108	protein coupled receptor 108
gRNA	guide RNA
hACE2	human Angiotensin-converting enzyme 2
HDR	homology directed repair
hGH	human growth hormone
HPV	Human papilloma virus
ILV	intraluminal vesicles
ITR	iverted terminal repeat
kb	kilobase pairs
LBD	ligand-binding domain
LE	late endosome
lssDNA	long single-stranded DNA
Luc/luc	luciferase
MAAP	Membrane Associated Accessory Protein
MCK	muscle creatine kinase
MHC II	MHC class II molecules
MOI	multiplicity of infection
MRN	Mre11/Rad50/NBS1 complex
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MVB	multivesicular body
MVB	multivesicular bodies
NLS	nuclear localization signal
OBD	origin binding domain
OHT	4-hydroxytamoxifen
ORF	open-reading frame
p53	Tumor protein P53
PBS	phosphate buffer saline
PEG	polyethylen glycol
PKA	protein kinase A
PKD	polycystic kidney disease repeat domains

PKR	dsRNA-activated kinase
PLA2	phospholipase A2
PLP	Myelin proteolipid protein
rAAV	recombinant adeno-associated virus
RAC1	Ras-related C3 botulinum toxin substrate 1
RBS	Rep-binding site
Rep	AAV gene coding for replication proteins
RNAi	RNA interference
RNP	ribonucleoprotein
scAAV	self-complementary AAV genome
shRNA	short hairpin RNA
siRNA	small interfering RNA
SKD1	Suppressor of K+ transport growth defect 1
ssAAV	single-stranded AAV genome
ssDBP	single-stranded binding protein
ssDNA	single-stranded DNA
SSR	site-specific recombinase
TGN	trans-Golgi network
TSG101	Tumor susceptibility gene 101
VA RNA	viral associated RNA
VP	viral protein
VRC	viral replication compartment
VSV-G	vesicular stomatitis virus glycoprotein G
WPRE	woodchuck hepatitis virus posttranscriptional regulatory element
wtAAV	wild-type adeno associated virus
YY1	Yin Yang 1 transcription factor

Abstract (EN)

Recombinant adeno-associated virus (rAAV) is an emerging vector with broad applications in basic research, particularly due to its non-integrative nature, robust long-term expression, and low immunogenicity. This dissertation presents several innovative advancements in the application and production of recombinant adeno-associated virus (rAAV) for transgenic practices, addressing key challenges in model development, allele conversion, and scalable production. The research aims to expand rAAV's potential in genetic research, particularly in creating mouse models and optimizing gene delivery systems. A major achievement was the development of a rapid, responsive mouse model for SARS-CoV-2 research, crucial during the COVID-19 pandemic. This model facilitated efficient testing of bi-specific antibodies and provided a valuable comparison to other transgenic models for viral infections. Additionally, the dissertation introduces a streamlined method for direct allele conversion in zygotes using *ex vivo* rAAV treatment, reducing breeding time and genotyping efforts. This approach supports the 3R principles of animal research by minimizing the number of animals required. Furthermore, a novel rAAV production method was developed, utilizing extracellular vesicles (EVs) as a biological matrix for scalable co-isolation of rAAV. This EV-AAV (EV-associated AAV) method offers a more efficient and less labor-intensive production process, maintaining or surpassing the performance of traditional techniques. Overall, this work contributes to the field of rAAV-based transgenic research, offering innovative methods that enhance the efficiency, scalability, and ethical standards of genetic research, with potential applications in both basic science and clinical therapy.

Abstrakt (CZ)

Rekombinantní adeno-asociovaný virus (rAAV) je vektor s širokou aplikací v základním a klinickém výzkumu, zejména díky své neintegrační povaze, robustní dlouhodobé expresi a nízké imunogenicitě. Tato disertační práce představuje několik inovativních přístupů v aplikaci a produkci rAAV pro transgenní praxi, přičemž řeší klíčové výzvy v oblasti vývoje modelů, konverze alel a škálovatelné produkce. Cílem práce je rozšířit potenciál rAAV v genetickém výzkumu, zejména při tvorbě myších modelů a optimalizaci systémů transferu genů. Hlavním úspěchem bylo vyvinutí rychlého, humanizovaného myšího modelu pro výzkum SARS-CoV-2, který byl klíčový během pandemie COVID-19. Tento model umožnil efektivní testování bispecifických protilátek a poskytl cenné srovnání s jinými transgenními modely pro výzkum SARS-CoV-2 infekcí. Disertační práce rovněž představuje zjednodušenou metodu přímé konverze alel v zygotách *ex vivo* pomocí rAAV, což vede ke zkrácení doby chovu a počtu potřebných zvířat pro dosažení cílového genotypu, což je v souladu s principy 3R v oblasti výzkumu na zvířatech. Dále byla v rámci této práce vyvinuta nová metoda produkce rAAV, která využívá extracelulárních vesikulů (EV) jako biologické matrice. Metoda spočívá v produkci složeného vektoru EV-AAV (EV-asociovaný AAV) a nabízí efektivnější a méně náročný výrobní proces rAAV, který si zachovává nebo dokonce překonává výkonnost tradičních izolačních technik. Celkově tato práce přispívá k vývoji transgenního výzkumu založeném na rAAV, přičemž nabízí inovativní metody, které zvyšují efektivitu, škálovatelnost a etické standardy genetického výzkumu s potenciálními aplikacemi jak v základním výzkumu, tak potenciálně v klinické terapii.

Table of Contents

1. Introduction

1.1 Adeno-Associated Virus

Wild-type adeno-associated virus (wtAAV) is a small, non-enveloped virus belonging to the *Dependoparvovirus* genus of the *Parvoviridae* family. It is classified as a dependovirus because it relies with its lytic cycle on helper viruses, such as adenovirus, herpesvirus or papillomavirus, for a productive replication[1,2].

In its natural state, wild-type AAV is not known to cause any human diseases and is considered non-pathogenic. However, it can establish latent infections in human cells, where its genome persists either as episomal DNA or integrates into the host chromosome mostly at specific site on chromosome 19 (19q13-qter), termed the AAVS1 site[3–5]. This unique characteristic of wild-type AAV has contributed to its use as a platform for developing recombinant AAV vectors (rAAV) suitable for gene therapy applications. Recombinant AAV vectors exhibit reduced toxicities and are non-pathogenic, making them safe and effective candidates for delivering therapeutic genes to target cells[6].

1.1.1 General molecular principle of AAV replication

The main AAV helper viruses are adenoviruses (i.e. AdV5), herpesviruses (i.e. HSV-1 or human cytomegalovirus HCMV) and papillomaviruses (i.e. human papillomavirus type 16, HPV-16), where the most studied interactions are with adenovirus type 5 (AdV5) and herpes simplex virus type 1 (HSV-1)[7–9]. Since helper-free production system is derived from AdV (specifically AdV5) helper virus, the following text will focus on the role this helper virus in wtAAV life cycle. The requirement for wtAAV/rAAV replication mediated by AdV is co-infection of the host cell with both viruses. In general, AdV types carry genes essential for AAV replication, namely E1A, E1B (E1B19K and E1B55K), E2A, E4orf6 and virus-associated RNA (VA RNA)[9]. In the following sections, the function of these factor in the context of AAV replication will be briefly described.

The genome of wtAAV is a single-stranded DNA molecule approximately 4.7 kilobases (kb) in size, containing two open reading frames (ORFs) that encode essential replication (Rep), capsid (Cap) and accessory proteins. Rep proteins (Rep78, 68, 52 and 40) are mediating the viral DNA replication in host cell, while Cap proteins (viral protein 1-3, VP1-3) form the viral capsid that encloses the viral genome. Accessory proteins, Assembly-Activating Protein (AAP), Membrane Associated Accessory Protein (MAAP) and protein X ensure capsid assembly, stabilization and release (Fig.1)[10].

AAV genes

AAV replication is dependent on the Rep proteins encoded by the **Rep genes**, which include Rep78, Rep68, Rep40, and Rep52. These proteins play a crucial role in facilitating DNA replication and encapsidation into the AAV capsid. Rep proteins contain an SF3 helicase domain, while the larger Rep proteins (Rep78 and Rep68) also possess an origin binding domain (OBD) at the N-terminus and a protein kinase binding site at the C-terminus. The N-terminal OBD domain of the large Rep proteins exhibits ATP-dependent endonuclease activity and binds specifically to the Rep-binding site (RBS) motif within the ITR (inverted terminal repeat) and p5 promoter sequence[11,12]. The C-terminal domain consists of a zinc-finger motif and a PKA inhibitor-like motif; the former is responsible for cell cycle arrest, while the latter inhibits AdV DNA replication[13–15]. This C-terminal domain is characteristic of Rep52 and Rep78, but not Rep40 or Rep68.

The Rep proteins display significant plasticity in their interactions with DNA substrates, allowing them to adopt different oligomeric states depending on the DNA they bind. Their activity and binding capacity are further regulated by post-translational modifications, such as phosphorylation, ubiquitination, and sumoylation, which can lead to either inactivation or degradation of the proteins[12,13].

Rep78 is an unstructured protein encoded by the only non-spliced transcript from the p5 promoter (Fig.1). All Rep proteins inhibit transcription from the p5 and p19 promoters in the absence of a helper virus. However, when the helper virus is present, this inhibition is relieved, allowing for the production of Rep78 and Rep68 proteins. The more active these promoters become, the more Rep78 and Rep68 are produced[16,17]. High levels of Rep78 result in the formation of hexameric oligomers, which exhibit ATPase, helicase, endonuclease, ligase, and DNA-binding activities[18]. These functions are essential for AAV DNA replication, integration, and gene regulation. Rep78 also interacts with various cellular proteins, significantly affecting the cell cycle by disrupting chromatin structure, inhibiting Cdc25A, sequestering E2F-1, activating pRb, and overall promoting S-phase arrest[14,19,20]. Furthermore, Rep78 interacts with several other proteins, including PC4, p53, and high-mobility group non-histone protein 1, while also blocking PKA and PrKK, and activating ATM and caspase-3. These interactions often lead to programmed cell death, making the abundance of Rep78 ultimately toxic for the host cell (Fig.2)[14,18].

Rep68 is a product of a C-terminally spliced mRNA transcript from the p5 promoter (Fig.1). It has endonuclease and DNA-binding activity, both mediated by its N-terminal domain. Rep68 is essential for the oligomerization of the Rep78/Rep68 complex, which is responsible for unwinding and processing viral DNA[12,13,21]. The C-terminal region of Rep68 contains an atypical 14-3-3 binding motif. When phosphorylated, Rep68 interacts with 14-3-3 proteins, which are regulators of cell division and apoptosis (Fig.2). This interaction causes the re-localization of Rep68 and reduces its binding to the ITR, ultimately leading to decreased viral DNA replication. The Rep68-14-3-3 interaction may regulate Rep68 levels and

modulate its role in the interplay between AAV and the helper virus[22]. In the absence of a helper virus, Rep68 contributes to the inhibition of the p5 promoter. In the presence and active replication, it is involved in arresting both the G1 and G2 phases of the cell cycle (Fig.2)[19].

Rep40 and Rep52 are products of spliced transcripts expressed from the p19 promoter (Fig.1). Both contain an SF3 helicase domain, Rep40 lacks the PKA binding motif, but includes a 14-3-3 binding motif. Similar to Rep52, the primary role of Rep40 is to facilitate the translocation of single-stranded DNA (ssDNA) into the AAV capsid (Fig.2)[22].

Rep52 exhibits 3'-5' helicase polarity, while Rep40 interacts with both the 3' and 5' ends of DNA, suggesting that Rep40 can bind both ssDNA and double-stranded DNA (dsDNA). This indicates that Rep40 likely unwinds dsDNA, converting it into ssDNA, which is more suitable for processing by Rep52[11]. The helicase activity of Rep40 is ATP-dependent and requires the protein to oligomerize. In contrast, Rep52 can bind and process DNA in its monomeric state[11,23].

Viral proteins (VPs) are products of the Cap gene, which is driven by the p40 promoter (Fig.1). This gene produces three transcripts: VP1, VP2, and VP3. Full-length Cap transcription generates VP1, while alternative splicing and start codons lead to the production of VP2 and VP3. These VP proteins are critical components of the AAV capsid, protecting the single-stranded DNA (ssDNA) genome from external factors[24,25].

VP1 is essential for efficient host cell infection and contains a phospholipase A2 (PLA2) domain. This domain remains enclosed within the capsid and is exposed to the surface only when the virus is endocytosed into the endosome. The enzymatic activity of PLA2 is crucial for endosomal escape and the subsequent transport of the viral genome into the nucleus (Fig.5)[26–28].

Both **VP1 and VP2** proteins contain a highly conserved Ser/Thr triplet motif, which is necessary for intracellular trafficking, transduction, viral assembly, and regulation of AAV genome expression[26,29,30]. Although VP2 is non-essential for viral infectivity, it is frequently used in capsid engineering and modification[31].

The most critical capsid protein is **VP3**, which forms the structural core of the AAV capsid. VP3 consists of eight β-barrel motifs, an α-helix, and loops that connect the β-barrels. While the motifs are highly conserved, the loops contain exposed variable regions that determine the binding affinity of different AAV serotypes to their specific target receptors[31–33].

Membrane assembly-activating protein (MAAP) is encoded within the VP1 region and is controlled by the p40 promoter (Fig.1). MAAP contains both non-polar and polar domains, which are essential for

membrane anchoring and mediating the extracellular secretion of AAV particles (Fig.8). Additionally, MAAP plays a protective role by preventing capsids from entering degradative subcellular compartments[34].

Assembly activating protein (AAP), like MAAP, is transcribed from the p40 promoter within the Cap gene region[35]. AAP is essential for AAV capsid assembly as it contributes to the stability and transport of VP proteins. The N-terminus of AAP contains hydrophobic domains that interact with the C-terminus of VP proteins, functioning as a chaperone[36]. Although the AAP-VP interaction is transient, it ensures the cotransport of VP proteins into the nucleus. In addition to its chaperone role, the C-terminal domain of AAP mediates the oligomerization of AAP itself. These oligomers interact with VP proteins, stabilizing and protecting them from degradation[35,37,38].

Protein X is a 155-amino acid auxiliary protein transcribed from the p81 promoter (Fig.1). It has been reported to enhance DNA replication, particularly in HEK293 cells during rAAV production[39].

The investigation into the molecular principles underlying coinfection with AdVs and AAVs led to the identification of essential AdV factors necessary for controlled rAAV production in cellular systems. The analysis and application of crucial AdV genes resulted in the development of helper-virus-free production systems, which are widely utilized for rAAV production in gene therapy research[40]. This innovation has significantly advanced the use of rAAV vectors in preclinical and clinical studies, offering promising prospects for the treatment of various genetic disorders and diseases[41].

AAV2 ITR (inverted terminal repeat) serves as the origin of replication and includes two smaller palindromes (B-B' and C-C') nested within a larger stem palindrome (A-A', Fig.2). The ITR can exist in two configurations: flip (with B-B' near the 3' end) and flop (with C-C' near the 3' end). A unique sequence, D, is found only once at each genome end, remaining single-stranded. The ITR contains a Rep-binding element (RBE) where Rep78 and Rep68 proteins bind (Fig.2)[42]. The RBE has a consensus sequence of 5'-GNGC-3'. Rep78 and Rep68, as ATP-dependent helicases, remodel the A-A' region to form a stem-loop, exposing the terminal resolution site (trs) in a single-stranded form[43,44]. These Rep proteins introduce a nick at the trs (Fig.2). An additional RBE (RBE') helps stabilize the interaction between the Rep proteins and the ITR[44].

Adenovirus genes

E1A and **E2A** are the most essential proteins for activation of the AAV p5 and p19 promotors and regulation of AAV Rep proteins[45,46] (Fig.1 and 2). Specifically, p5 promoter contains E1A-inducible element, which is activated upon E1A binding, and in parallel E1A interacts with Yin Yang 1 (YY1) ubiquitous transcription factor (originating from the host cell) repressing p5 promoter. The E1A-YY1 interaction

neutralizes the repression effect of YY1[47,48]. Due to their size E1 genes are often present in helper-free systems in form of a constitutively expressed cassette integrated in producer cells genome (Fig.3)[41].
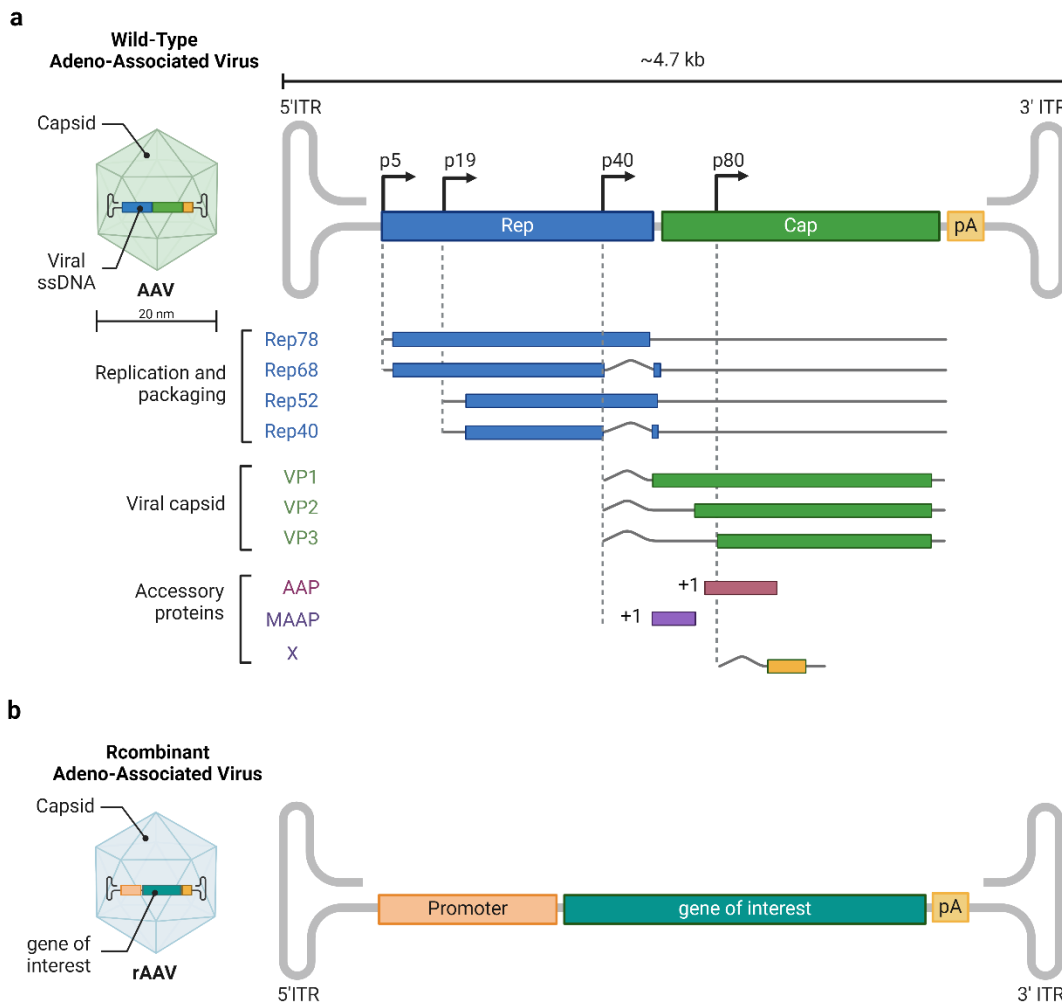

Figure 1: Schematic comparison of wtAAV and rAAV genome
(a) Wild-type AAV genome consists of genes coding for replication protein Rep78, Rep68 (expressed by p5 promoter), Rep52 and Rep40 (expressed by p19). Downstream capsid and accessory proteins (AAP and MAAP) are expressed from p40 promoter. Within Cap gene resided gene coding for protein X controlled by p80 promoter. (b) rAAV genome has Rep and Cap genes replaced by transgenic cassette of interest. Rep and Cap genes are supplied in trans as pAAV-Rep/Cap plasmid during rAAV production. Both wtAAV and rAAV genomes are flanked by two ITR sequences (adapted from Catalán-Tatjer *et al.* 2024).
Abbreviations: ITR - inverted terminal repeat, AAP – assembly activating protein, membrane assembly-activating protein – MAAP. Created in Biorender.

E2A is a single-stranded DNA binding protein (ssDBP) apart from mediating AAV DNA replication, the protein is also involved in processing of Rep proteins (Fig.2). Rep protein can further interact with E2A

and recruit it to AAV viral replication compartments (VRC[1*]). E2A protein plays role in variety of processes from AAV replication, mRNA processing and export, or capsid production[46,49,50].

E1B gene complex codes for two oncoproteins **E1B19K** and **E1B55K** with distinct functions. E1B19K subunit mediates inhibition of E1A-induced apoptosis and triggers autophagy during AdV infection[51]. E1B19K factor is not required for AAV replication, but rather increases AAV titer yields[52]. E1B55K binds and inhibits the tumour suppressor protein p53, thus promotes cell-cycle progression and inhibit apoptosis[53]. In the context of AAV replication E1B55K functions together with E4orf6 (Open reading frame 6 of early region 4) as a helper factor (Fig.2). This complex promotes AAV second strand synthesis and viral DNA replication (Fig.2)[54]. In addition, E1B55K enhances export of AAV mRNA and inhibits export of host cell mRNA, both increasing AAV gene expression[54].

E4orf6 has a major role as part of E1B55K-E4orf6 complex in mRNA export and AAV replication, specifically in second-strand synthesis of the viral genome (Fig.2). However, most important role of E4ofr6 is promotion of degradation of Mre1, crucial part of Mre11/Rad50/NBS1 (MRN) complex, an important mediator of the cellular DNA-damage response[55]. MRN complex represses AAV transduction and replication. It has been reported that E4orf6 is the most crucial, AdV-origin factor, needed for effective AAV production. Rep and Cap proteins can compensate for lack of other abovementioned factors in AAV production, but they cannot compensate for E4orf6. However, E4orf6 also negatively impacts AAV replication by promoting the degradation of newly assembled AAV capsids and Rep53 via ubiquitin-dependent manner. In contrast, Rep proteins can inhibit expression of E4ofr6 gene to establish balance between two separate stages of AAV replication. This ambiguous function of E4orf6 is an example of hurdle between AdV and AAV replication. This interaction allows for fine-tuning of expression profiles of both viruses in coinfected cells[56,57].

VA RNA (viral associated RNA) is generally responsible for RNA-mediated inhibition of the cellular innate immune protein dsRNA-activated kinase (PKR). Inhibition of PKR prevents induction of PKR-mediated shut-down of general translation, leading to efficient virus protein synthesis and interferon resistance of AdVs[58]. AAV p40-generated RNAs naturally trigger PKR in the host cells, which often leads to inhibition of AAV protein synthesis. Therefore, VA RNA enhances expression of AAV Cap proteins by inhibiting PKR pathway (Fig.2)[59].

[1*] VRC: Viral replication compartments are specialized structures in host cells where viral genomes and replication machinery concentrate, enhancing viral replication and protecting against host defenses.
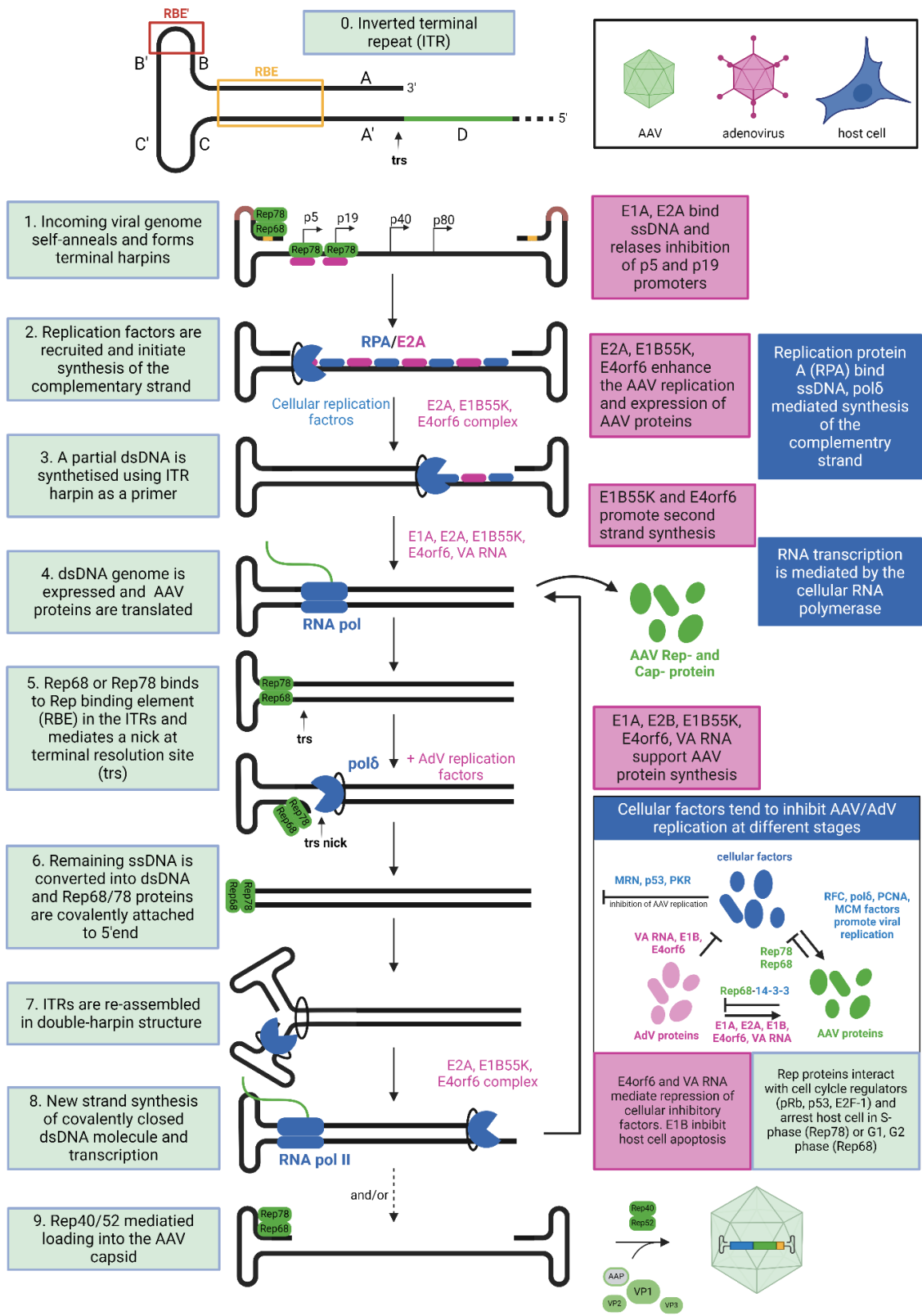
RBE'

0. Inverted terminal repeat (ITR)

B'   B

RBE

A
3'

C'   C

A'   D   5'

trs

AAV   adenovirus   host cell

1. Incoming viral genome self-anneals and forms terminal harpins

Rep78 Rep68   p5   p19   p40   p80

Rep78 Rep78

E1A, E2A bind ssDNA and releses inhibition of p5 and p19 promoters

2. Replication factors are recruited and initiate synthesis of the complementary strand

RPA/E2A

Cellular replication factros

E2A, E1B55K, E4orf6 complex

E2A, E1B55K, E4orf6 enhance the AAV replication and expression of AAV proteins

Replication protein A (RPA) bind ssDNA, polδ mediated synthesis of the complementry strand

3. A partial dsDNA is synthetised using ITR harpin as a primer

E1B55K and E4orf6 promote second strand synthesis

4. dsDNA genome is expressed and AAV proteins are translated

E1A, E2A, E1B55K, E4orf6, VA RNA

RNA pol

AAV Rep- and Cap- protein

RNA transcription is mediated by the cellular RNA polymerase

5. Rep68 or Rep78 binds to Rep binding element (RBE) in the ITRs and mediates a nick at terminal resolution site (trs)

Rep78 Rep68

trs

polδ   + AdV replication factors

Rep78 Rep68

trs nick

E1A, E2B, E1B55K, E4orf6, VA RNA support AAV protein synthesis

6. Remaining ssDNA is converted into dsDNA and Rep68/78 proteins are covalently attached to 5'end

Rep78 Rep68

Cellular factors tend to inhibit AAV/AdV replication at different stages

cellular factors

MRN, p53, PKR

inhibition of AAV replication

RFC, polδ, PCNA, MCM factors promote viral replication

VA RNA, E1B, E4orf6

Rep78 Rep68

Rep68-14-3-3

7. ITRs are re-assembled in double-harpin structure

AdV proteins

E1A, E2A, E1B, E4orf6, VA RNA

AAV proteins

8. New strand synthesis of covalently closed dsDNA molecule and transcription

E2A, E1B55K, E4orf6 complex

RNA pol II

E4orf6 and VA RNA mediate repression of cellular inhibitory factors. E1B inibit host cell apoptosis

Rep proteins interact with cell cylcle regulators (pRb, p53, E2F-1) and arrest host cell in S-phase (Rep78) or G1, G2 phase (Rep68)

and/or

9. Rep40/52 mediatied loading into the AAV capsid

Rep78 Rep68

Rep40 Rep52

AAP   VP1

VP2   VP3

Figure 2: Contribution of host cell, AdV and AAV factors to AAV replication
Schematic description of the AAV genome replication shows structure and elements of AAV inverted terminal repeat (ITR,0.). Green boxes describe roles of AAV derived molecules (green shapes) in AAV genome (black) replication. Purple boxes and text describe function and effects of AdV-derived on AAV genome replication. Blue text, shapes and boxes depict host cell originated molecules involved in AAV genome replication (adapted and modified from Meier *et al.* 2020 and Catalán-Tatjer *et al.* 2024).
Abbreviations: RBE - Rep-binding element; RNA pol II – RNA polymerase II; pol δ – DNA polymerase delta, MRN - Mre11/Rad50/NBS1 (MRN) complex; PKR - dsRNA-activated kinase; RFC – replication factor C; PCNA - Proliferating cell nuclear antigen; MCM – mini-chromosome maintenance protein; pRb - Retinoblastoma protein; p53 – tumor suppressor protein; E2F-1 - Transcription factor E2F1. Created in Biorender.

1.1.2   Helper-Virus Free System and Production of rAAV vector

The life cycle of AAV relies on helper viruses. The helper virus-free system has been derived from AdV5 interaction with AAV in HEK293 cells. This system employs selected functional AdV factors—namely E1 (E1A and E1B), E2A, E4orf6, and VA RNA—to provide the necessary helper functions.

The system consists of three main components:

1. a transfer plasmid carrying the gene of interest.
2. two packaging plasmids: Rep/Cap and Helper.
3. a transgenic HEK293 cell line.

In this setup, the E1A and E1B factors are constitutively expressed by the transgenic HEK293 cells. The helper plasmid integrates E2A, E4orf6, and VA RNA. The pAAV-Rep/Cap plasmid supplies Rep genes (Rep78, Rep68, Rep40), capsid genes (VP1-3), and accessory genes (AAP, MAAP).

In the helper virus-free method, the three plasmids are simultaneously transfected into HEK293 cells. Within the transfer plasmid (pAAV-GOI), the therapeutic expression cassette is flanked by AAV inverted terminal repeats (ITRs), which serve as the AAV replication origin and enable selective packaging of the transgenic cassette into AAV capsids (Fig.3)[41].

Following transient transfection in mammalian cells, mature AAV virions are harvested from the cell culture medium or cell lysate and subsequently purified for downstream purposes. The development of the adenoviral helper (pAd-Helper) plasmid has facilitated this process, making it independent of viral infection (Fig.3)[41,60].

AAV vectors produced using this method have been extensively utilized in preclinical studies to establish proof of principle and have also been used to generate clinical-grade AAV vectors for human trials[40,61,62].
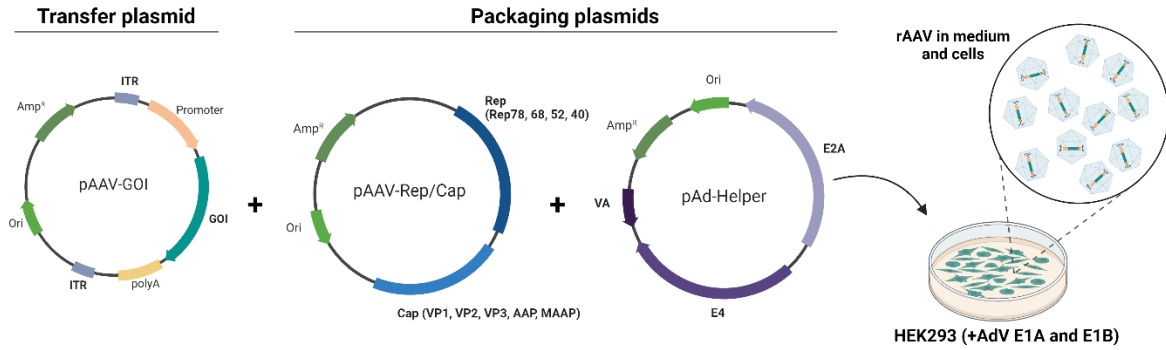
Figure 3: Key components of helper virus-free rAAV production
Transfer plasmid (pAAV-GOI) carries gene of interest (GOI) flanked by ITR sequence; Packaging plasmid: replication/capsid plasmid (pAAV-Rep/Cap) supplies Rep78, Rep68, Rep52 and Rep40, VP1-3, MAAP, AAP, X proteins enhancing replication of rAAV and packaging in host cell. Helper plasmid (pAd-Helper) activates rAAV replication via AdV5-derived genes: E2A, E4, VA RNA gene. Packaging plasmid and transfer plasmid are delivered in HEK293 producer via triple transfection using a transfection reagent (e.i. polyethylenimine – PEI or calcium phosphate – $CaPO_4$). rAAV are harvested from culture medium and cell lysate. Created in Biorender.

1.1.3 rAAV serotypes and tropism

The names of AAV serotypes include two numerical designations separated by a slash, indicating the ITR and serotype. For instance, AAV2/8 has a type 2 ITR (genome element) and a type 8 serotype (capsid element). Most vectors contain type 2 ITRs derived from the AAV2 serotype. The serotype refers to the type of capsid used to package an rAAV vector, determining the virus's ability to infect specific cell types[63].

AAV serotypes are defined by their lack of efficient cross-reactivity with neutralizing sera specific to other known serotypes. Based on this definition, only AAV1–5 and AAV7–9 are considered true serotypes. Variants like AAV6, 10, and 11 do not fit this definition[64–66]. AAV6's serology is nearly identical to AAV1, and the serological profiles of AAV10 and AAV11 are not well characterized [67].

The AAV capsid surface exhibits affinity to primary and secondary receptors that facilitate binding to target cells. The variable strength of interaction with these receptors defines the AAV serotype, with each serotype exhibiting specific tropism and a preference for infecting specific cell populations. Researchers can select the most suitable AAV serotype for targeting the desired tissue by choosing the appropriate virus serotype, dictated by variations in capsid proteins (Fig.4)[6]. Capsid proteins are the subject of thorough study in the context of developing tissue- or cell-type-specific rAAV particles, or conversely, broadening their tropism. The focus of the study is on naturally occurring variants of capsid proteins or unique variants generated through machine-guided design to obtain capsids with unique delivery features[68–70]. Fig.4 summarizes commonly used capsid serotypes and their tropism towards certain tissue and cell types[71–73]. While many serotypes overlap in their affinity, some serotypes are more efficient at delivering to certain tissues than

others. More detailed information on cell membrane surface factors interacting with specific capsid serotypes, can be found in sub-chapter 1.1.4 rAAV cargo delivery and expression of GOI, Tab. 1.
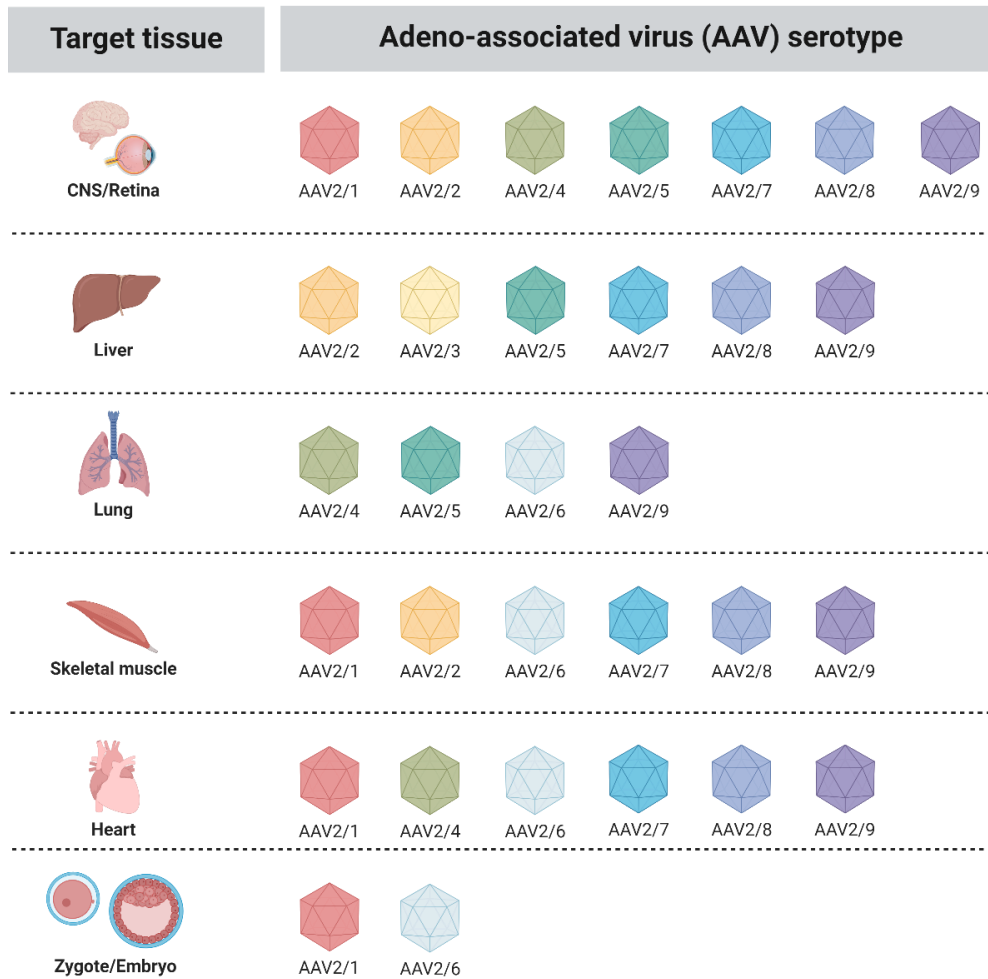

Figure 4. Adeno-associated virus serotypes and their tropism
Overview of tropism of the main AAV serotypes by definition of Earley *et al.* 2020. Data compiled from Wu *et al.* (2006) and Issa *et al.* (2023). Serotypes capable of transducing zygotes/embryos are as reported by Chen *et al.* (2019). Created in Biorender.

1.1.4   rAAV cargo delivery and expression of GOI

The recombinant adeno-associated virus (rAAV) life cycle involves several critical steps: attachment, internalization, intracellular trafficking, nuclear transport, uncoating, and gene expression.

The life cycle begins with the attachment of the virus to the target cell (Fig.5). During this step, the rAAV capsid interacts with a primary (glycan) receptor on the cell surface, while integrins and growth factor receptors serve as co-receptors (secondary protein receptor) to facilitate viral entry. For instance, the AAV2 serotype primarily binds to heparan sulphate proteoglycan and utilizes co-receptors such as the fibroblast

growth factor receptor[74], hepatocyte growth factor receptor[75], laminin receptor[76], and integrins αVβ5 and α5β1[77,78]. These interactions enable the AAV2 serotype to exhibit broad tissue tropism (Tab.1)[79].

Other AAV serotypes interact with different proteins and receptors on the cell membrane. For example, AAV4 binds to O-linked 2,3-sialic acid with affinity for the central nervous system (CNS), specifically targeting ependymal cells.[79] AAV5 interacts with N-linked sialic acid and uses the platelet-derived growth factor receptor as a co-receptor, facilitating broad interaction with CNS cells and airway epithelial cells[79]. AAV1 and AAV6 also interact with N-linked sialic acid, targeting tissues such as embryonic cells[73]. Additionally, AAV9 binds to N-linked galactose, enabling it to cross the blood-brain barrier[80,81].

Genome-wide screens have identified novel host proteins essential for rAAV transduction, termed entry factors. One example is KIAA0319L, also known as AAV receptor (AAVR), a glycoprotein that has been proven to act as a universal AAV receptor, mediating the transduction of both wild-type and engineered rAAV capsids[82]. Another key entry factor is G protein coupled receptor 108 (GPR108), a member of the G protein-coupled receptor (GPCR) superfamily, which facilitates many interaction events across various AAV serotypes (Tab.1)[83].

Table 1: The major receptors, co-receptors, entry factors, involved in the interaction of various AAV serotypes with their respective cell surface receptors.

Serotype	Glycan (primary) receptors	Protein (secondary) co-receptors	Entry factor	Reference
AAV2/1	AAVR, SIA	unknown	AAVR, GPR108	84,85
AAV2/2	AAVR, HSPG	integrin αVβ5 and α5β1, hFGFR, hHGFR, LamR	AAVR, GPR108	74,75,77,78,84,86
AAV2/3	AAVR, HSPG	hHGFR, LamR	AAVR, GPR108	76,84,87
AAV2/4	SIA	unknown	GPR108	
AAV2/5	AAVR, SIA	PDGFR	AAVR	84,88
AAV2/6	AAVR, SIA	EGFR	AAVR, GPR108	84,85,89
AAV2/7	unknown	unknown	AAVR, GPR108	
AAV2/8	AAVR	LamR	AAVR, GPR108	76,84
AAV2/9	N-linked galactose of SIA, AAVR	LamR	AAVR, GPR108	76,81,84

Platelet-derived growth factor receptor (PDGFR), Hepatocyte growth factor receptor (hHGFR), Fibroblast growth factor receptor (hFGFR), Laminin receptor (LamR), Epidermal growth factor receptor (EGFR), α2–3 and α2–6 N-linked sialic acid (SIA), universal AAV receptor (AAVR/ KIAA0319L)

Internalization

Following receptor binding, the virus is internalized by the cell through endocytosis. It has been reported that approximately 13% of AAV2 particles in the vicinity of HeLa cells are internalized within 1.2 seconds. Additionally, recombinant AAV2 (rAAV2) particles are trafficked to the nucleus within 15 minutes of internalization, highlighting the rapid and highly conserved nature of this process[90,91].

The process of internalization can occur through several pathways, including caveolin- and clathrin-dependent mechanisms, as well as clathrin- and caveolin-independent pathways[92]. One major clathrin- and caveolin-independent route is micropinocytosis, mediated by the Ras-related C3 botulinum toxin substrate 1 (RAC1) protein, along with clathrin-independent carriers (CLICs) in glycosylphosphatidylinositol (GPI)-anchored protein-enriched endosomal compartments (GEECs, Fig.5)[93,94]. Interestingly, it has been reported that clathrin, caveolin, or micropinocytosis may not be essential for rAAV transduction. Studies investigating specific internalization pathways typically inhibit these pathways to assess their impact on rAAV transduction efficiency. These studies have demonstrated that such inhibition often has no effect on rAAV uptake or, in some cases, can even enhance uptake in certain cell types, while reducing it in others[94,95]. Another important factor influencing rAAV transduction is co-infection with adenovirus, which can modulate the cellular processes involved in rAAV internalization[96]. Thus, the internalization pathway utilized by rAAV is both cell type- and context-dependent[95].

Cellular trafficking

Once inside the cell, the virus must be transported to the nucleus through a process known as intracellular trafficking, which involves sorting the endocytosed cargo. To reach the nucleus, rAAV must navigate several trafficking compartments, including the early endosome (EE), late endosome (LE), and the Golgi apparatus. Viral trafficking is initiated in the EE, which communicates with the trans-Golgi network (TGN) through bidirectional exchange. Efficient transduction of rAAV particles often requires transit through the Golgi (Fig.5) [97–99]. Some serotypes, such as rAAV2, can undergo endocytic processing through the LE, although the maturation of rAAV in the LE may be cell type-dependent (Fig.5)[100]. A key factor affecting rAAV transduction is the degradation of viral capsids through the ubiquitin-proteasome system, autophagy, or endoplasmic reticulum-associated protein degradation (ERAD). Inhibiting the ERAD or the ubiquitin-proteasome system has been shown to increase the transduction efficiency of several rAAV serotypes. Additionally, mutating one or more surface-exposed tyrosine (Tyr) or threonine (Thr) residues prevents their phosphorylation, thereby reducing capsid ubiquitination[101]. This, in turn, lowers the degradation rate and enhances the likelihood of the capsid reaching the nucleus. The activation of the ubiquitin-proteasome system, ERAD, or autophagy is often dose-dependent, with higher doses of rAAV more likely to trigger these responses. Notably, proteasomal degradation of rAAV capsids can provoke a capsid-specific T-cell

response, which is a common issue in clinical applications[102]. Proteasome inhibitors have been shown to reduce this immune response[103].

Endosomal escape

To enter the nucleus, rAAV must first escape from endosomal or trans-Golgi network (TGN) compartments into the cytoplasm. Once released, the capsids accumulate in the perinuclear space[104]. The process of endosomal escape is facilitated by conformational changes in response to acidic pH, such as the exposure of the phospholipase A2 (PLA2) domain within the VP1 protein. Other contributing factors to this escape process include the self-proteolytic activity of the rAAV capsid, the entry factor GPR108, and the protease activity of cathepsins (Fig.5)[105–107].

Although the role of autoproteolysis remains poorly understood, the role of GPR108 has been identified as crucial for both endosomal escape and nuclear entry, functioning as a key nuclear factor. Interestingly, the role of GPR108 appears to be dependent on the PLA2 domain[83]. However, rAAV5 serotype lacks the PLA2 domain, and instead, it interacts via the ectodomain of PKD1, one of two immunoglobulin-like polycystic kidney disease (PKD) repeat domains (PKD1 and PKD2) in the AAVR receptor[108].

Following endosomal escape, rAAV capsids accumulate near the perinuclear space and are transported into the nucleus. Both VP1 and VP2 proteins contain four basic regions (BR1-4) that are involved in protein interactions, sorting, and signal transduction within the host cell. Importantly, these regions also harbour nuclear localization signal (NLS)-like sequences[109,110]. The β-importins recognize the NLS sequences in VP1 and VP2, mediating the nuclear import of the capsids through nuclear pore complexes (NPC, Fig.5)[111].

Notably, studies have reported that only 17% of rAAV particles present in the cytoplasm successfully traverse into the nucleus, highlighting the potential for increasing transduction efficiency by enhancing nuclear import mechanisms[112].

Genome release and expression

In the nucleus, the rAAV genome is released from the capsid, although complete disassembly of the capsid is not required for genome release[113]. Studies using fluorescence microscopy, subcellular fractionation, and 3D transduction analysis have shown that rAAV2 capsids can be sequestered in the nucleolus, enabling secondary transduction. This theory suggests that intact capsids accumulate within the nucleolus, where they are later released and disassembled in the nucleoplasm (Fig.5)[114].

In the absence of Rep and Cap genes, as well as a helper virus, rAAVs cannot replicate but can still transcribe their genomes. As with wtAAV replication, the inverted terminal repeats (ITRs) in the rAAV genome enable the priming for complementary strand synthesis. The single-stranded AAV (ssAAV)

genome must be converted into a double-stranded DNA (dsDNA) form to beging the transcription. This can occur either through the synthesis of a complementary strand or by hybridization between complementary strands. However, in the absence of helper factors, the synthesis of the second strand is significantly reduced in rAAV vectors, making complementary strand alignment a more common method for producing a transcriptionally active DNA molecule[115,116]. Self-complementary rAAV (scAAV) genomes overcome this limitation through mutated ITRs, which facilitate molecular recombination, leading to concatemerization and circularization of the viral genome. Notably, ssAAV and scAAV genomes exhibit distinct release kinetics, with the release rate inversely correlated to the length of both ssDNA and scDNA genomes (Fig.5)[113].

Once released, rAAV genomes tend to persist in postmitotic cells mainly as episomal concatemers. However, a small fraction of rAAV genomes can integrate into the host genome, i.e. AAVS1 site. Despite the low integration rate and reduced risk of genotoxicity, the integrative potential of rAAV remains an important focus of study due to safety concerns in clinical applications[117].
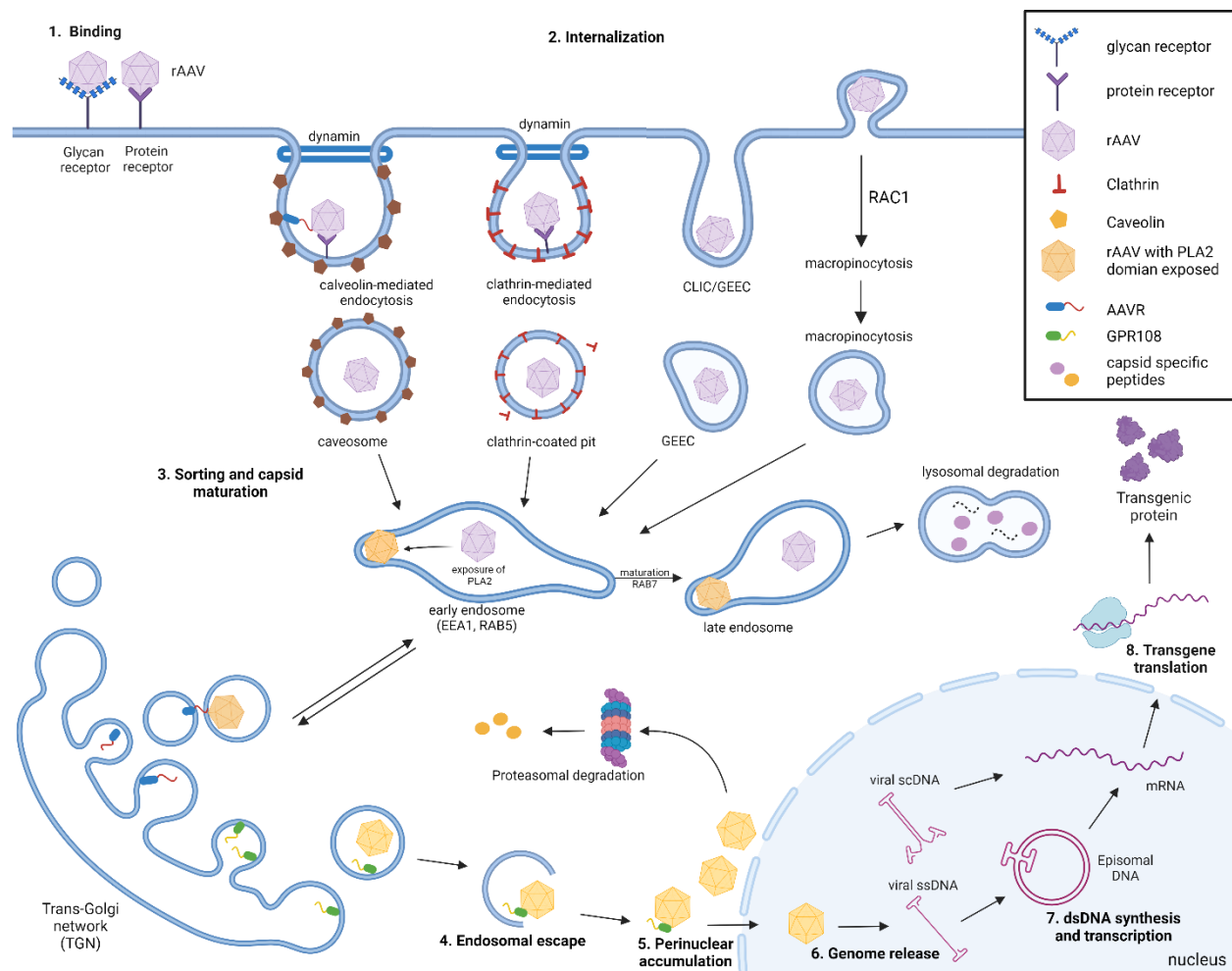
Figure 5. Mechanism of rAAV entry and transduction
1.Binding: rAAV attachment to the cell surface is mediated by primary glycan receptors and stabilized by secondary protein receptors. The specific receptors required depend on the rAAV serotype and cell type. The role of entry factors such as KIAA0319L and GPR108 in facilitating rAAV entry at the cell surface remains to be fully understood, however they are sequestered and recycled in TGN. 2.Internalization: rAAV enters the cell through multiple endocytic routes, including clathrin- and caveolin-mediated endocytosis, macropinocytosis, and clathrin-independent pathways, such as those involving clathrin-independent carriers (CLICs). 3.Sorting and Capsid Maturation: Once internalized, the rAAV capsid undergoes conformational changes in sorting compartments like endosomes and the Golgi. Key entry factors, such as KIAA0319L and GPR108, are primarily localized in the Golgi and play a role in capsid maturation. Capsids unable to mature are translocated into lysosome for degradation. 4.Endosomal Escape: Efficient cytoplasmic escape and transduction depend on exposure of the phospholipase A2 (PLA2) domain at the N-terminus of the capsid. Following escape from endosomes, vector particles accumulate near the perinuclear space. 5.Perinuclear Accumulation: rAAV particles gather around the perinuclear region and enter the nucleus via nuclear pore complexes, a process mediated by importins like importin-β. At this stage, capsids can be hyper ubiquitinated and degraded in proteasome. Genome Release: Inside the nucleus, the rAAV genome is released from the capsid (6.). 7.dsDNA synthesis and Transcription: The single-stranded AAV (ssAAV) genome is converted into double-stranded DNA, or in the case of self-complementary AAV (scAAV), the genome is already double-stranded, enabling transcription of the transgene. 8.Transgene Translation: The transgene is subsequently translated into the target protein, resulting in successful gene expression in the host cell.
Adapted from Dhungel *et al.* (2021). Abbreviations: dsDNA, double-stranded DNA; scAAV, self-complementary AAV; ssAAV, single-stranded AAV; CLIC, clathrin-independent carrier; GEEC, glycosylphosphatidylinositol-anchored protein (GPI-AP)-enriched compartment; TGN, trans-golgi network. Created in Biorender.

1.2 Clinical potential of rAAV vectors and their production

Recombinant AAV vectors have emerged as a versatile and promising tool for gene therapy, offering efficient and stable transduction of a wide range of target tissues, long-lasting gene expression without integration into the host genome, and minimal immunogenicity. Their favourable safety profile, combined with a broad tissue tropism encompassing 16 distinct types, makes them suitable candidates for treating a variety of rare genetic disorders or neurodegenerative diseases[62,118–120].

The identification of essential elements for AAV replication has enabled the development of production methods using various animal cell-based platforms, eliminating the need for helper viruses and improving the safety and control of rAAV manufacturing[37,62,120]. These advancements have contributed to the successful development of marketed gene therapy products, such as LUXTURNA® (Novartis) for Leber congenital amaurosis, ZOLGENSMA® (Novartis) for spinal muscular atrophy, HEMGENIX® (CSL Behring) for haemophilia B, ROCTAVIAN® (BioMarin) for haemophilia A, ELEVIDYS® (Sarepta Therapeutics) for Duchenne muscular dystrophy, and BEQVEZ® (Pfizer) for haemophilia B. The clinical application of rAAV vectors holds great potential for delivering safe and effective gene therapies to address unmet medical needs and provide transformative treatments for patients.

The listed gene therapy products often face market demand challenges due to their high price, which limits access for potential patients. For example, ZOLGENSMA costs €1.9 million per course of treatment. This high price is largely due to the complexity of the manufacturing process and stringent quality control measures. For instance, the recommended dose of ZOLGENSMA is 1.1×10^{14} vector genomes (vg) per

kilogram of body weight, amounting to 2.2×10^{15} vg for a child weighing 20 kg. Each current dose of ZOLGENSMA (5.5 or 8.3 mL per vial) contains 2×10^{13} vg per mL, requiring a total of 14 vials per treatment[121]. While optimized rAAV production and purification methods can yield high concentrations of clinical-grade virus, the costs remain high due to the intricate process required to optimize the highest yield with a favorable full-empty capsid ratio. This purification step can significantly reduce the final titer, increasing the demand for input material[62].

Current rAAV production methods involve either the transfection of specific plasmids or infection with viruses carrying the required genes, both of which pose scalability challenges and limitations in capsid loading efficiency[62].

Human embryonic kidney 293 (HEK293) cells are widely used for the production of recombinant proteins and rAAV in clinical applications[122–124]. Their popularity stems from their human origin, which supports proper post-translational modifications, their well-understood metabolism *in vitro*, and their prior approval by both the FDA and EMA for biopharmaceutical production[38,123,124]. For instance, ZOLGENSMA is manufactured using transgenic HEK293 cells that contain essential adenoviral genes. The remaining necessary components for rAAV production are provided via triple plasmid transfection (Fig.3)[40]. Alternatively, there are also platforms that allow supply of the essential components through viral infection or by stable expression in transgenic cell lines.

One example of viral infection-based production is the baculovirus expression vector system (BEVS) using Sf9 insect cells. Sf9 cells are capable of performing post-translational modifications crucial for the bioactivity and stability of therapeutic proteins[125]. Through genetic modification, insect-specific α(1,3)-fucosylation can be eliminated[126]. In this system, Sf9 cells are infected with three recombinant baculoviruses: one delivering the Rep78/52 genes, another expressing Cap proteins, and a third providing the transfer plasmid. BEVS offers an advantage of no need for exogenous helper genes since Sf9 cells naturally contain all required helper functions[127,128]. While this platform supports large-scale rAAV production, it faces challenges such as a low full-to-empty capsid ratio and difficulties in removing baculovirus from final preparations[129]. Despite these challenges, the BEVS platform has been successfully employed in producing therapeutics for human papillomavirus (HPV) - CERVARIX, and gene therapies like HEMGENIX for hemophilia B[130,131].

1.3 Application of AAVs in research and disease modelling

rAAV vectors have emerged as powerful tools for disease modelling due to their ability to efficiently and safely deliver genes into target cells. By incorporating disease-causing mutations or genes associated with specific disorders into AAV vectors, researchers can create *in vitro* and *in vivo* models that recapitulate key

aspects of human diseases. These models enable the investigation of disease mechanisms, identification of therapeutic targets, and evaluation of potential treatments. Additionally, AAV-mediated gene delivery allows for precise control over gene expression levels and spatiotemporal patterns, facilitating the study of disease progression and therapeutic interventions in a controlled experimental setting. Disease modelling using rAAV vectors holds great promise for advancing our understanding of various genetic and acquired disorders, ultimately paving the way for the development of novel therapies.

1.3.1   ssDNA delivery to zygotes and embryos *ex vivo*

In context of transgenic practice, single-stranded DNA (ssDNA) homology templates are demonstrated to be more efficient for integration into the genome through homology-directed repair (HDR), particularly highlighted with long ssDNA (lssDNA) homology templates in the context of the EASI (efficient additions with ssDNA inserts)-CRISPR method[132]. However, synthesizing complex lssDNA can pose challenges, and its efficiency may vary depending on the specific cassette of interest. Furthermore, in some instances, lssDNA can adopt rigid secondary structures, rendering them inaccessible to HDR machinery. It's important to note that lssDNA templates can typically only be delivered using invasive methods, such as microinjection or electroporation, to ensure their entry into target cells (Fig.6)[132,133].

The life cycle of adeno-associated virus (AAV) involves the packaging of either sense or antisense ssDNA into a virion. Consequently, an AAV virion comprises a single-stranded DNA genome and capsid proteins (VP1, VP2, VP3). Capsid assembly occurs through interactions between VP proteins and the internal terminal repeat (ITR) sequences of the AAV genome. If the genome remains within the packaging size limit of approximately 5 kbp, the resulting assembly yields a biologically active virion[6,134,135].

The production of recombinant AAV (rAAV) vectors circumvents common challenges associated with lssDNA synthesis. rAAV is often produced in a controlled replication cell line, ensuring efficient replication, packaging, and stabilization of the ssDNA genome, followed by non-invasive delivery into target cells. These properties render the rAAV vector suitable for delivering larger transgenic constructs into a mammalian genome. This capability was demonstrated by Chen *et al.* through the development of a method termed CRISPR-READI[73].

CRISPR-READI involves delivering ssDNA templates in the form of rAAV into zygotes, leveraging the natural properties of AAV serotype 1 and 6 capsids to traverse the zona pellucida and subsequently interact with moieties on the zygote's cell membrane. Upon uptake by the zygote, the viral particle undergoes uncoating, and the ssDNA genome is translocated into the nucleus to serve as a homology template. This approach has proven highly efficient compared to standard methods involving pronuclear injections of plasmids or linear double-stranded DNA (dsDNA) templates. The CRISPR-READI procedure entails brief

treatment of zygotes with rAAV carrying the HDR template, electroporation of Cas9/gRNA ribonucleoprotein complex (RNP), and overnight incubation with HDR-AAV[136].

The primary advantages of this method include reduced physical stress, facilitated membrane crossing by the virus through a natural biological mechanism, and the reliable delivery of an efficient dose of the HDR template. Additionally, electroporation of RNPs ensures the effective delivery of a precise amount of site-specific nuclease. These factors collectively contribute to the high efficiency of HDR template insertion at the target locus, with success rates of up to 60% observed among all founder animals. However, it's important to acknowledge potential challenges associated with the use of rAAV vectors, such as concatemerization of the transgenic cassette at the target locus, as highlighted by Suchy *et al.*[137]. Therefore, rigorous evaluation and validation of the generated models are essential to ensure the accuracy and integrity of genetic modifications achieved through the CRISPR-READI approach. This cautious approach will help maximize the method's potential for precise and reliable genome editing in diverse biological applications.

Comparing the EASI-CRISPR and CRISPR-READI methods, both approaches represent significant advancements in the field of genome editing utilizing ssDNA templates. EASI-CRISPR is based on delivery of "naked" long single-stranded DNA (lssDNA) molecule to the pronucleus. The method highlights the efficacy of lssDNA homology templates in facilitating precise genome modifications through HDR. However, challenges in the synthesis, purification, stability, and delivery of these templates can limit its broader adoption (as illustrated in Fig. 6). In contrast, CRISPR-READI optimizes ssDNA delivery by utilizing rAAV vectors. This method harnesses the natural properties of rAAV to stabilize the ssDNA genome, ensuring efficient cellular uptake and HDR template insertion.As a result, CRISPR-READI offers streamlined and effective genome editing, demonstrating reduced physical stress on embryos and achieving high efficiency in target locus insertion. Ultimately, while EASI-CRISPR underscores the potential of ssDNA templates, CRISPR-READI provides a practical and efficient approach by integrating ssDNA delivery within the versatile AAV vector system, paving the way for transformative applications in precision genetic engineering. Methodological comparison of the two method is depicted in Fig.6.
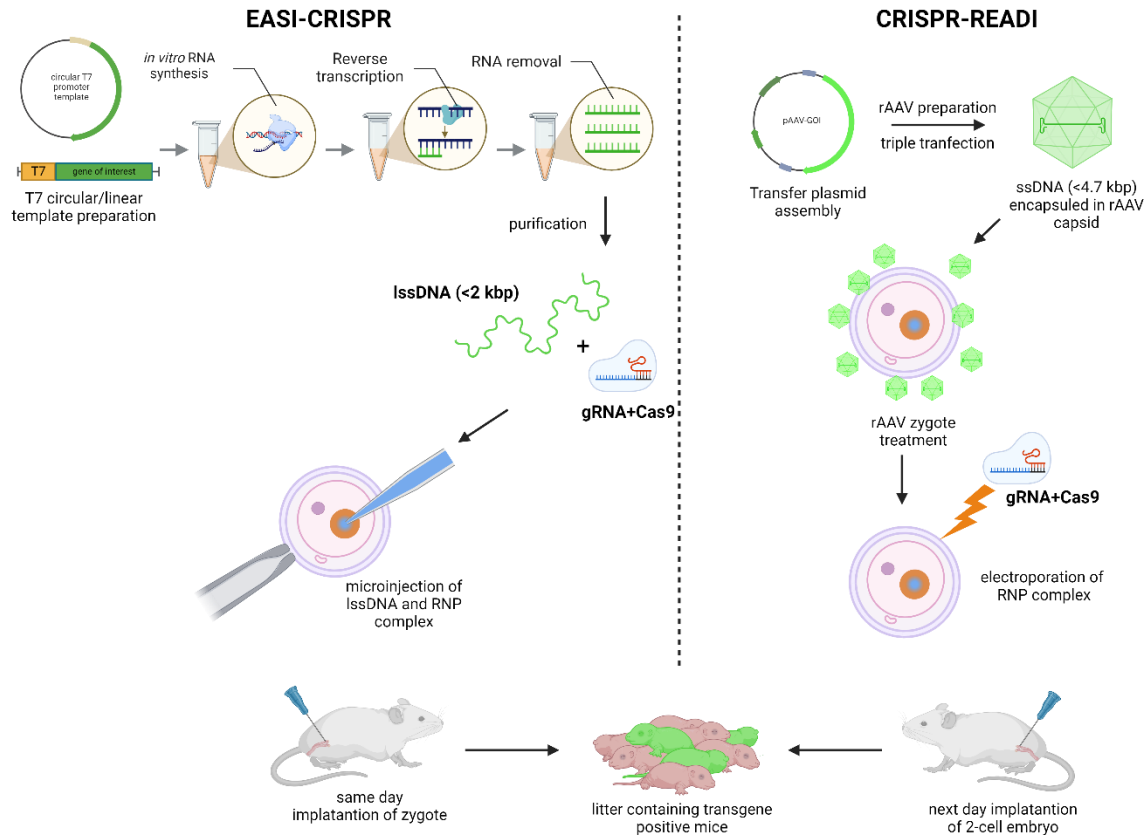
Figure 6: Comparison of EASI-CRISPR and CRISPR-READI.
EASI-CRISPR requires synthesis of long-single EASI-CRISPR requires the synthesis of long single-stranded DNA (lssDNA) via in vitro RNA transcription, followed by reverse transcription into cDNA. The RNA is then removed, and the cDNA is purified. The lssDNA synthesis can be also performed chemically or via asymmetric PCR. The purified lssDNA (up to 2 kb) is used for pronuclear injection (PNI) in combination with the Cas9-gRNA ribonucleoprotein complex (RNP). The treated zygote is implanted into the oviduct of a surrogate female on the day of treatment. The surrogate female gives birth to genetically altered pups resulting from the PNI delivery of lssDNA and RNP. The CRISPR-READI method involves the assembly of a transfer plasmid carrying the gene of interest, which is used for rAAV production. The resulting virus packages a single-stranded AAV genome up to 4.7 kb (ssAAV/ssDNA). The viral genome is delivered into the zygote in a non-invasive manner using the zygote's natural primary and secondary receptors. After incubation of the zygote with the virus, RNP is delivered via electroporation. The zygote is incubated with the virus overnight, and the following day, the 2-cell embryo is implanted into the surrogate female. The surrogate female gives birth to genetically altered pups based on viral delivery of ssDNA and RNP electroporation. Figure was compiled based on Miura *el al.* (2018) and Chen *et al.* (2019). Created in Biorender.

1.3.2   Application of rAAV in basic research

Recombinant AAVs (rAAVs) are a powerful tool for deciphering biological processes and the origins of diseases in model organisms. They are widely employed to deliver various biologically active molecules, including reporters[138], gene coding sequences (CDS)[139], recombinases[140], and genome editors[141,142]. Efficient, tissue-specific delivery of these molecules can have a profound impact, enhancing our understanding of complex biological mechanisms or disease pathology.

For instance, the application of reporter system in rAAV has resulted in development of RNA interference (RNAi)-based method for dynamic regulation of the expression of therapeutic transgenes delivered by adeno-associated virus vectors. This approach uses small interfering RNA (siRNA) or short hairpin RNA (shRNA) to suppress transgene expression and employs a synthetic oligonucleotide, REVERSIR, to reverse this suppression. Luc (luciferase) reporters were utilized in the study to monitor and quantify transgene expression *in vivo*, enabling precise evaluation of the effectiveness of RNAi-mediated regulation. This method allows for on-demand control of transgene dosage, providing valuable insights into transgene expression and regulation. The system has potential clinical applications, addressing challenges such as variable transgene expression and dose-dependent toxicity in AAV-mediated gene therapy. The study presents specific siRNA sequences that regulate transgene expression while minimizing off-target effects, offering a promising framework for dynamic control in gene therapy[138].

The article by Piepho *et al.* (2023) is an example of reduced CDS of a large gene packaged in rAAV[139]. The study investigates the long-term cardiac efficacy of AAV-mediated micro-dystrophin gene therapy in a severe Duchenne muscular dystrophy (DMD) model. The serotype AAV9 was used, which is known for its efficient muscle targeting, particularly in skeletal and cardiac tissues. The MHCK7 promoter, a hybrid promoter derived from muscle creatine kinase (MCK) and α-myosin heavy chain (α-MHC) promoters, was employed to ensure muscle-specific expression of micro-dystrophin. The therapy resulted in sustained cardiac improvements, including enhanced ejection fraction and reduced fibrosis, suggesting potential for long-term cardiac preservation in DMD. The study showed that after the administration of the AAV9 vector, treated animals exhibited normalized cardiac function, confirmed by echocardiographic and histological analyses, with reduced signs of heart failure, structural damage, and improved heart morphology over time. Furthermore, micro-dystrophin expression was stably maintained, demonstrating that this gene therapy could provide long-lasting cardiac benefits in DMD patients[139].

The cell-type-specific recombination in the murine central nervous system (CNS) has been highlighted in the study focused on using rAAV vectors with the astrocyte-specific mGfaABC1D (recombinant glial fibrillary acidic protein promoter) promoter, combined with PHP.eB (recombinant AAV serotype) capsid, to achieve astrocyte-specific recombination. The researchers addressed a major challenge of nonspecific transduction by dividing the rAAV system into two vectors: one expressing FlpO recombinase and the other expressing Cre recombinase flanked by Flp recognition sites. This two-vector system, combined with the astrocyte-tropic AAV-F capsid, greatly improved the specificity and efficiency of recombination in astrocytes, potentially reducing off-target effects[140].

rAAV vectors enable efficient delivery of CRISPR-based editors. Liao *et al.* has presented a novel CRISPR/Cas9 system for gene activation that avoids DNA double-strand breaks [141]. The system utilizes a

catalytically inactive Cas9 (dCas9) to recruit transcriptional activators (MS2-p65-HSF1) at specific gene loci, inducing epigenetic changes without permanent DNA modifications. Delivered through AAV2/9 vectors, the system was tested *in vivo* across various mouse models, demonstrating its therapeutic potential. In Duchenne muscular dystrophy, acute kidney injury, and type I diabetes models, target gene activation improved muscle strength, kidney function, and insulin production, respectively. By using modified dead guide RNAs (dgRNAs) to avoid DSBs, the system offers a safer alternative to traditional CRISPR editing. The rAAV9 vectors facilitated efficient delivery across tissues, and multiplexed gene activation was also demonstrated. This study highlights the system's potential for safe, *in vivo* gene activation and epigenetic therapy for human diseases[141].

Another example of the remarkable application of AAV vector is Wolter *et al.* work. The article discusses the development of a novel CRISPR system using *Staphylococcus aureus* Cas9 (SaCas9) delivered via AAV2/9 to target the long non-coding RNA Ube3a-ATS, which silences the paternal allele of UBE3A. The research focuses on Angelman syndrome, a neurodevelopmental disorder caused by the loss of maternal UBE3A. This CRISPR system, delivered *in utero*, successfully reactivates the paternal UBE3A allele in neurons. The treatment showed long-term gene activation in Angelman syndrome mouse models, improving both behavioral and anatomical deficits without significant off-target effects[142].

An important aspect common to all of the aforementioned studies is the emphasis on tissue specificity and the effects of delivered molecules in the target tissue. As noted, AAV serotypes exhibit varying tropisms for specific tissues. Another factor that enhances specificity is the use of tissue-specific promoters to drive gene expression. Therefore, a strategic combination of an AAV serotype and a specific promoter can significantly improve both the specificity and efficiency of gene therapy, tailoring the treatment to the desired target tissue for optimal results.

1.3.3 Generation of humanized mouse model

Despite their primary application in delivering specific expression cassette for therapeutic purposes, recombinant adeno-associated virus (rAAV) vectors can also be utilized to generate unique humanized models. The principle of humanization involves delivering a human orthologue into a wild-type animal to express a human protein, thereby humanizing the animal or specific tissues of interest.

Standard production of a transgenic mouse model typically involves several steps: creating a transgenic construct, delivering and integrating this construct into the genome, selecting founder animals, confirming germ-line transmission of the gene of interest (GOI) through breeding, and expanding the selected line to obtain a sufficient number of animals for experimentation. This process can take a minimum of nine months to complete.

In contrast, transgene delivery using rAAV offers a faster and more scalable method to generate humanized models without the complexities of genome integration. rAAV allows for the delivery of expression systems up to 5 kbp in a tissue-specific manner, depending on factors such as the rAAV serotype used and the tissue-specific promoter incorporated into the vector[143,144].

Developing a humanized model with rAAV involves packaging the desired expression construct into the vector and administering the vector at a specific titer into the animal. The entire process of vector development and production typically lasts up to two months. Once a sufficient amount of virus is obtained, it can be directly applied to a cohort of animals, bypassing the need for extensive breeding and lineage assessment associated with traditional transgenic models. This approach using rAAV streamlines the generation of humanized models, making it a rapid and efficient alternative to traditional transgenic methods for modelling human diseases or studying specific gene functions *in vivo*[143–145].

The potential of rAAV-mediated humanization was demonstrated during the COVID-19 pandemic, which was characterized by a shortage of relevant mouse models for SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) research, specifically humanized ACE2 (hACE2) mouse models. The mouse ACE2 protein is not recognized by the SARS-CoV-2 virus, thus it cannot bind and facilitate viral entry into target cells. To overcome this limitation, expression of hACE2 in target cells or tissues is necessary to enable SARS-CoV-2 entry and replication[146,147].

The unavailability of transgenic hACE2 models during the pandemic prompted the development of various new transgenic models and alternative methods for rapid humanization. These methods predominantly rely on viral vectors, among which recombinant adeno-associated virus (rAAV) plays a significant role. rAAV vectors offer an efficient means to introduce the hACE2 gene into mouse models, enabling the creation of humanized mice capable of supporting SARS-CoV-2 infection and replication. This approach allowed researchers to swiftly generate relevant models for studying COVID-19 pathogenesis and evaluating potential therapies, demonstrating the versatility and utility of rAAV-mediated humanization during critical biomedical research endeavours [143,144,147].

1.4 Delivery limitations

The wild-type adeno-associated virus (wtAAV) genome is approximately 4.7 kb in size, which closely aligns with the packaging limit for a fully functional recombinant AAV (rAAV) genome. Although there have been reports indicating that rAAV vectors can carry genomes ranging from 4.7 to 8.9 kb, vectors exceeding a genome size of 5.2 kb generally exhibit significantly reduced transduction and expression efficiency[134,148]. This packaging limit is consistent across all AAV serotypes; however, some serotypes demonstrate a higher tolerance for larger genomes compared to others. For instance, Allocca *et al.* reported

that the AAV2/5 serotype can package genomes up to 8.9 kb, but subsequent efforts to standardize the packaging of such large genomes were unsuccessful[134,148,149]. The efficiency of transduction with a vector carrying a genome larger than 5.2 kb is influenced by factors such as titer yield, multiplicity of infection (MOI), and genome stability within the host cell [134,149].

The packaging process of oversized AAV genomes is challenging, often resulting in the production of empty capsids or capsids carrying truncated genomes. These factors contribute to a decreased final titer, which can be up to tenfold lower than that of vectors carrying normal-sized genomes (4.7 kb)[134,150].

Oversized genomes are frequently truncated at the 5' end, necessitating compensation by another genome that contains the complete 5' end. Through homologous recombination, these truncated genomes can assemble a complete genome within the host cell, thereby recovering the potential for gene expression. This intracellular vector recombination is enhanced by a higher multiplicity of infection (MOI) when using vectors carrying oversized genomes, in contrast to the MOI typically required for normal-sized AAV[134,151].

AAV capsids enter the cell via endocytosis and are processed through intracellular trafficking via the late endosome. During this trafficking process, the AAV capsid undergoes conformational changes that expose the VP1 N-terminus, making it susceptible to ubiquitination in the cytoplasm (Fig.5)[152]. Ubiquitination primarily occurs during the endosomal escape of virions into the cytoplasm and their subsequent transport to the nucleus. Oversized genomes often fail to efficiently escape the late endosome, leading to their degradation in the lysosome. Alternatively, they may become hyper-ubiquitinated and are subsequently eliminated by the proteasome (Fig.5)[149].

Circumventing the AAV packaging limit is a key challenge in gene therapy research. Recombinant AAVs are highly regarded as vectors for gene therapy due to their low immunotoxicity and ability to provide long-term, robust gene expression[118]. However, their limited packaging capacity makes them unsuitable for therapies that require the delivery of large gene constructs. Overcoming this limitation is essential to fully unlocking the potential of rAAVs in a broader range of gene therapies.

Several strategies have been developed to overcome the AAV packaging limit. One such approach involves reducing the size of the target gene by creating "mini" versions, which are constructed by identifying and preserving only the essential functional domains of the gene (Fig.7a). This reduction is based on a detailed understanding of the gene's critical regions and their effective application, with some minigenes' functionality being predicted *in silico*[153–156]. The example is the Δ17–48 mini-dystrophin gene, discovered by England *et al.*, which was designed to retain the key functional domains of the dystrophin gene. This mini-dystrophin was subsequently optimized and adapted for clinical testing[153,154].

The recombination-based assembly of large transgenic constructs often utilizes two split vectors, each carrying a portion of the transgene cassette. Within the host cell, these partial genomes undergo homologous recombination to form a full-length coding sequence (Fig.7b). The efficiency of this strategy appears to depend on the characteristics of the host tissue, particularly whether it contains dividing or non-dividing cells[157]. Muscle tissue has been shown to be a suitable target for this approach due to the presence of proliferating myoblasts, as demonstrated in studies on Duchenne muscular dystrophy and similar dysferlinopathies[158–161]. In contrast, terminally differentiated tissues, such as photoreceptors, exhibit lower recombination rates compared to muscle tissue[162]. Homologous recombination is a key DNA repair pathway in rapidly dividing cells, making it more effective in proliferating myoblasts, which possess a robust DNA recombination machinery[163].

The trans-splicing and hybrid strategies exploit the propensity of AAV genomes to persist in the nucleus as episomes through concatamerization. This process is driven by the sequence homology of the inverted terminal repeats (ITRs) (Fig,7c). However, the formation of concatamers connected via ITR sequences can pose challenges for recombination-based assembly. Specifically, the presence of an ITR at the junction between two AAV genomes can interrupt the coding sequence, potentially causing a frameshift in the transgene of interest. To address this, donor and acceptor splice sites are strategically placed to facilitate the removal of the ITR-mediated junction through the intron splicing process (Fig.7c)[164–166].

To further improve the assembly of specific concatemers and, consequently, precise splicing, the homology-directed junction of the ITRs can be enhanced by introducing a homologous region shared between the two compatible AAV genomes. For example, a highly recombinogenic 77-base-pair homology sequence derived from the filamentous phage F1 has been shown to support the precise assembly of concatemers (Fig.7d)[162].

The capacity limit of AAV vectors can also be circumvented by leveraging distinct molecular interactions at the protein level, such as protein trans-splicing or AAV capsid modification. Protein trans-splicing involves the joining of two separate polypeptides through a peptide bond to form a complete protein. This process is facilitated by intervening protein segments known as inteins, which function similarly to self-cleaving introns in DNA. In a split system, the intein is divided into two segments (split-intein), with each segment fused to one half of the transgenic protein. When the split-inteins connect, they undergo self-excision, mediating the formation of a peptide bond between the two halves of the transgenic protein (Fig.7e)[167].

The successful application of the split-intein system has been demonstrated in several cases, including the assembly of Becker dystrophin in muscle, the retina-specific ABC transporter 4 (ABCA4), and Centrosomal Protein of 90 kDa (CEP90) in the retina[168,169]. A critical factor for the success of the split-intein system is

the precise and functional placement of the split-intein within the target protein, ensuring that the protein maintains its proper structure and function following trans-splicing[157].

An alternative strategy to increase the AAV packaging capacity involves modifying the capsid proteins to enhance their stability during the packaging of larger genomes, while simultaneously maintaining the genome in a more condensed state. This approach was demonstrated by Tiffany and Kay, who showed that altering the lumenal charge of the AAV capsid allowed for the successful packaging of transgenes up to 6.2 kb in size. This capsid modification improves the efficiency of packaging larger genomes without compromising the structural integrity of the vector[170].

All of the aforementioned strategies enable the delivery of transgenes exceeding the 4.7 kb packaging limit of AAV vectors. However, these approaches typically do not achieve the same packaging, transduction, and expression efficiency as standard-sized AAV vectors. This reduction in efficiency is influenced by several factors, including the characteristics of the host cells (whether they are dividing or non-dividing), the efficiency of recombination or trans-splicing events, and the ability to achieve uniform distribution and expression within the target tissue.
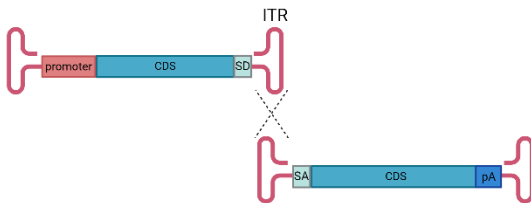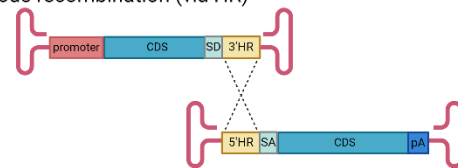

a) Mini-gene
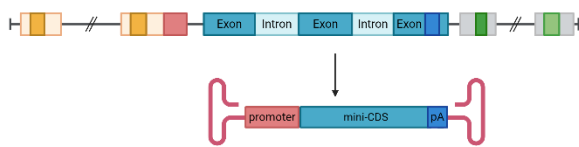
b) Homologous recombination mediated assembly (via HR)
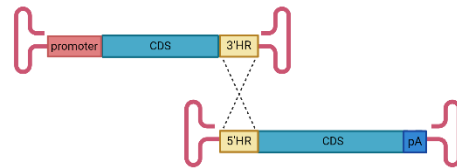
c) Concatemerization and trans-splicing

d) Hybrid: trans-splicing and concatemerization (via ITR) or homologous recombination (via HR)
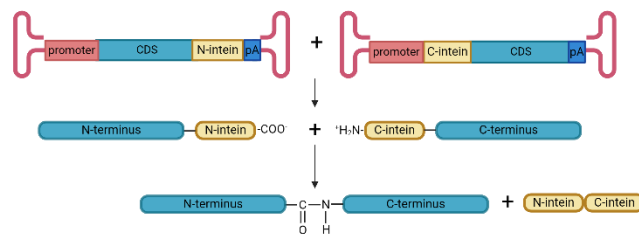
e) Protein trans-splicing (via inteins)

Figure 7. Approaches to circumventing rAAV capacity limit.
(a) The full-length transgene can be modified to retain only essential protein domains, creating a functional miniature gene that fits within a single AAV vector. (b) The dual-vector strategy, where the transgene is split into two parts that are delivered together and reassembled upon co-infection based on homologous regions. (c) The transgene fragments can recombine through concatamerization facilitated by the intervening ITRs, ITRs are further removed by splicing during mRNA transcription (d) Hybrid system based on homologous recombination and splicing, a hybrid approach uses homologous recombination via a highly recombinogenic homology sequence (HR), e.g. highly recombinogenic 77-base-pair homology sequence, HR is spliced out during transcription via SD and SA sites. (e) Protein trans-splicing method involves expressing transgene fragments as independent polypeptides that can be seamlessly joined by split-inteins system (Adapted from Marrone *et al.* 2022).
Abbreviations: ITR = Inverted Terminal Repeat, CDS = Coding Sequence, SD = Splice Donor, SA = Splice Acceptor, pA = Polyadenylation Tail, HR = Homologous Recombination sequence. Created in Biorender.

1.5   rAAV production advancements

Various methods have been employed for isolating recombinant rAAVs, including affinity resins, ion-exchange chromatography, and the standard approach of separating producer cell lysates using a discontinuous gradient followed by ultracentrifugation[171–173]. Although effective, this process is laborious, time-consuming, expensive, and posses challenges for achieving the desired quantity and quality of rAAVs.

The most commonly used isolation method is separation via a discontinuous gradient. This method enables the separation of rAAV particles based on their specific density from cell debris and organelles. The density of rAAV particles (ρ = 1.216 g/ml) corresponds to a 40% iodixanol solution, an iodine-containing density gradient medium. rAAV capsids are typically found in the 40% iodixanol fraction, effectively separating them from lower-density cellular components. After collecting this fraction, the iodixanol medium must be removed by ultrafiltration [174].

Ultrafiltration is critical for proper rAAV purification, as it ensures the removal of the density medium without significant loss of the functional titer, which can occur due to aggregation or precipitation[174]. Although iodixanol-based separation offers a favourable full-to-empty capsid ratio and high titers compared to scalable purification methods like ion-exchange or affinity chromatography, the main bottleneck remains the efficient removal of the density gradient medium without compromising rAAV yield[172–175].

1.5.1   Extracellular vesicles and Extracellular vesicles associated with AAVs

In recent years, a novel approach utilizing extracellular vesicles (EVs) alongside AAV vectors, termed EV-AAVs, has emerged. Initially proposed by Maguire *et al.*, EV-AAVs, or vexosomes, are microvesicle/exosome-associated AAV vectors derived from HEK293 producer cells using the standard AAV production method. EVs are isolated from the cell culture medium through differential centrifugation. This innovative strategy demonstrated that AAV capsids associate with both the surface (protein corona) and interior of microvesicles/exosomes, leading to enhanced transduction efficiency compared to free AAVs and increased resistance to neutralizing antibodies[176]. The association of AAV with EV particles can result from AAV infection and its subsequent internalization within the host cell's endosomal trafficking

system. In the case of AAV producer cells—those that facilitate AAV replication—mature capsids are exported from the nucleus and released either through host cell apoptosis or exosome-mediated secretion. Additionally, EVs may integrate into microvesicles during the cell membrane budding process (Fig. 8)[177–179]. The close interaction between AAV biogenesis and the endosomal system enables the isolation of EV-AAV composite particles[180].

Extracellular vesicles (EVs) are nano-sized proteolipid particles secreted by almost all cell types into the extracellular milieu. They represent a diverse population of nanoparticles, encompassing apoptotic bodies (AB, 50-5000 nm), membrane-derived microvesicles (50-1000 nm), and endosome-derived exosomes (50-150 nm, Fig.8)[181]. The size and composition of these EVs vary based on the specific biosynthetic pathways involved, the cell type, and the prevailing metabolic conditions. EVs are integral to intercellular communication in both normal physiological processes and pathological states[182]. Furthermore, their intrinsic ability to interact with cellular membranes and engage in diverse endogenous processing pathways positions them as a promising delivery system[181].

In recent years, exosomes, a specific subset of EVs, have emerged as a focal point of research interest. Exosomes originate within the endosomal system, specifically within late endosomes where intraluminal vesicles (ILVs) accumulate within multivesicular bodies (MVBs, Fig.8). Exosome formation unfolds in two stages. Initially, the endosomal membrane becomes enriched with tetraspanins like CD9 and CD63. Subsequently, cargo loading occurs via ESCRT-dependent or ESCRT-independent mechanisms. The ESCRT-dependent pathway involves recruiting proteins such as ESCRT 0-III, Alix, Tsg101, Chmp4, and SKD1, leading to the formation of MVBs containing CD63, MHC II, ubiquitinated proteins, and KFERQ-containing proteins. Conversely, ESCRT-independent pathways, mediated by proteins like Syntenin-1, Syndecan, and heparinase, also give rise to exosomes containing typical markers such as PLP, CD63, CD81, and TSG101[182,183]. These proteins serve as markers for exosomal biogenesis, cargo selection, and release[184].

Aside from their role in exosomal biogenesis, proteins like ALIX and TSG101 also participate in plasma membrane budding, a common feature in microvesicle biosynthesis. Microvesicles differ from exosomes as they are formed through the outward budding of the plasma membrane, involving distinct proteins and significant lipid and protein rearrangements between the inner and outer membrane layers (Fig.8)[184,185].

Apoptotic bodies (ABs) are generated during programmed cell death (apoptosis) through the externalization and blebbing of the plasma membrane. ABs contain fragments of cellular organelles and degradative enzymes such as proteases, DNases, and RNases (Fig.8)[184].

During the isolation of exosomes and microvesicles, it is imperative to separate them from cellular debris and apoptotic bodies to prevent degradation and obtain a well-defined particle population. Various

approaches, such as differential centrifugation based on size and density, are employed to achieve this. Despite their distinct protein compositions, exosomes and microvesicles can be co-isolated due to their similar physical characteristics. Removal of apoptotic bodies, which can induce degradation, is typically achieved through filtration or differential centrifugation steps at 10,000-20,000G[184].

Advancements in vexosome technology include modifying their composition and enhancing production within producer cells. Studies showed that pseudotyping vexosomes with the vesicular stomatitis virus glycoprotein G (VSV-G) broadened their tropism, resulting in increased yields of both viral particles and EVs. Overexpression of the tetraspanin CD9 in producer cells also enhanced exo-AAV production [176,186].

Based on reported data, EV-AAVs offer several advantages over traditional AAV vectors. They provide a combination of transient and stable vector characteristics, with EVs delivering active RNAs and proteins for a specific duration while AAVs carry a stable DNA genome[179]. EV-AAVs are rapidly and directly isolated, exhibit resistance to neutralizing antibodies, can be pseudotyped for tissue targeting or blood-brain barrier penetration, and offer versatile cargo delivery capabilities[187].
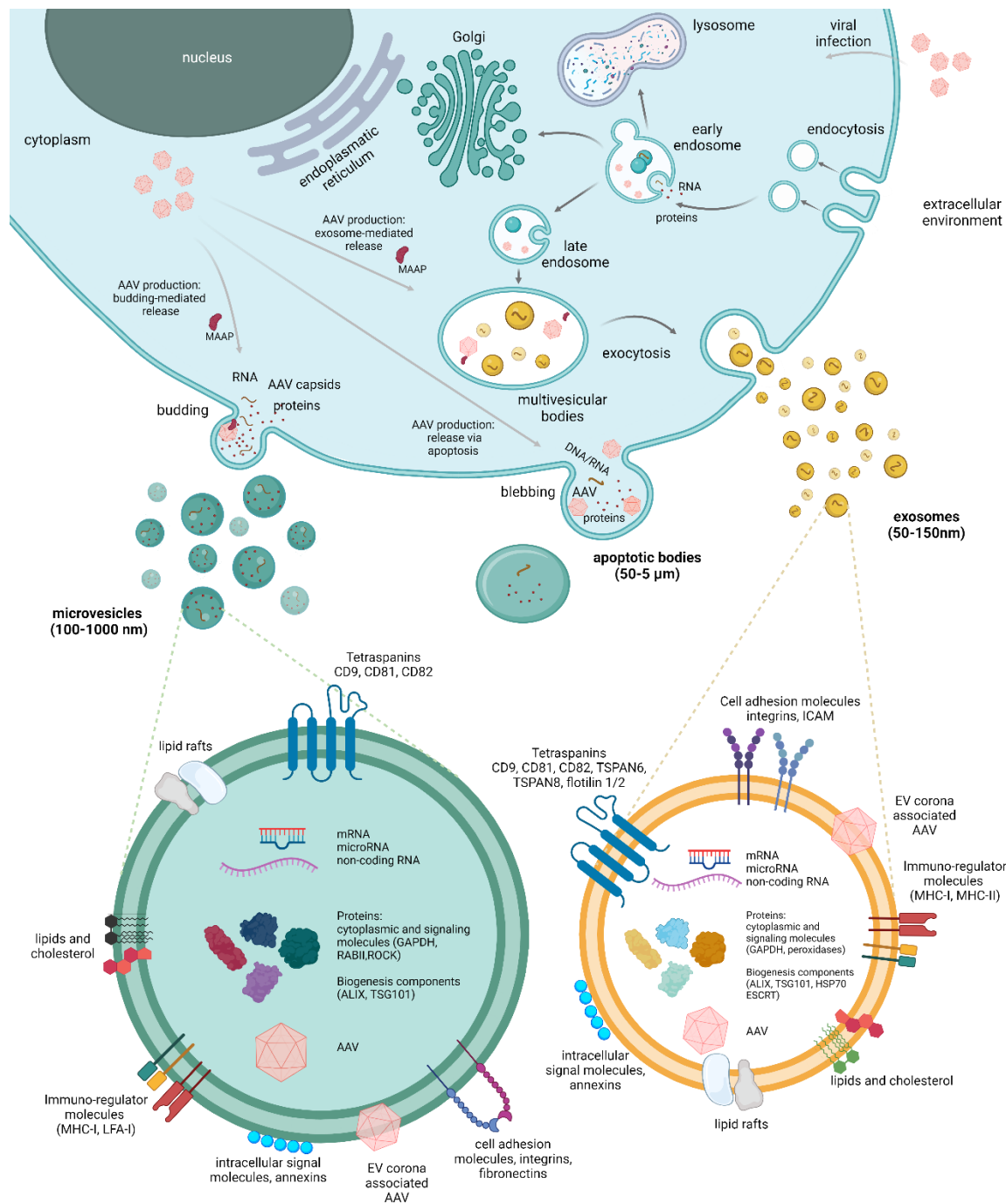
Figure 8. Biogenesis of extracellular vesicles and release of rAAV from producer cells.
Mammalian cells release various types of extracellular vesicles, including exosomes, microvesicles, and apoptotic bodies, each with distinct biogenesis pathways and lipid/protein compositions. Microvesicles and exosomes carry specific RNA and protein molecules, reflecting the cell type of origin or its physiological state. Apoptotic bodies contain DNA, RNA, and proteins from apoptotic cells, conveying signals of apoptosis. In producer cells, rAAV particles are exported from the nucleus and released from the host cell through apoptosis, either as free viruses or within apoptotic bodies. The regulated release of rAAV particles can occur through exosome-mediated exocytosis or controlled plasma membrane budding (microvesicles) with main involvement of accessory protein MAAP (membrane assembly-activating protein). This figure is based on observations by Deshetty *et al.* (2024), Zhang *et al.* (2018), Gyorgy and Maguire (2017), and Wang *et al.* (2024). Created in Biorender.

2. Aim of the study

This work aims to demonstrate a practical application of the rAAV system in various approaches, including disease modelling in adult mice, targeted genome engineering *ex vivo*, and exploring the potential role of extracellular vesicles in rAAV acquisition and application

Currently, recombinant adeno-associated virus vectors represent an emerging delivery system for gene therapy. rAAV vectors are characterized by robust, long-term expression, a non-integrative nature, and relatively low toxicity, making the system an ideal candidate for gene therapy delivery. However, therapeutic purposes may not be the sole application of this system; on the contrary, diverse uses can be derived from rAAV's properties. The system can be repurposed and utilized for *in vivo* disease modelling or transgenic disease model generation.

Aims:

Aim 1: To establish rAAV vectors for the rapid generation of transient *in vivo* mouse model.

Aim 2: Investigation of new rAAV applications in genome engineering.

Aim 3: Development of an alternative rAAV Purification through the strategic utilization of EVs.

3. Materials and methods

3.1 Materials and methods that are not included in published papers

Materials and methods that are not included in published papers are only those mentioned in manuscript (in preparation) "*Scalable Production of rAAV Vectors via Extracellular Vesicle-Mediated Purification for Gene Therapy and Transgenesis*"

3.2 List of methods that were included into published papers

- Subcloning of mRNA template plasmids and rAAV tranfer plasmids
- Production and purification of recombinant adeno-associated virus
- Titration of rAAV
- Production of mRNA *in vitro*
- Mouse zygote electroporation and pronuclear injection
- Genotyping of mouse tissue and blastocyst
- Tissue samples and embryo imaging
- Intranasal application of rAAV into mouse
- Construct validation in Neuro-2A cells

4. Results

The results section is divided into three key parts. Part 1 highlights the potential of rAAV in disease modelling, demonstrating how the strategic delivery of rAAV carrying human gene variants can transiently humanize mice, allowing for the rapid generation of disease models. Part 2 explores the innovative use of rAAV as a delivery vector for site-specific recombinases, facilitating precise allele conversion *ex vivo*. Part 3 presents the development of a scalable and efficient method for the production and purification of rAAV using extracellular vesicles.

4.1 Part 1a: Transient disease modelling *in vivo* using rAAV

The COVID-19 pandemic highlighted the urgent need for humanized mouse models carrying the human Angiotensin-converting enzyme 2 (hACE2) receptor. Mice are naturally resistant to SARS-CoV-2 infection due to the inability of the murine Ace2 protein to mediate viral entry. As a result, the development of humanized models became essential. Despite the existence of models such as K18-hACE2, HFH4-hACE2, and Ace2-hACE2, the pandemic led to limited availability, creating a significant demand for rapidly generated, functional, and accessible models. Traditional methods of mouse model generation, which can take up to two years, were insufficient to meet this urgent need.

In response, we utilized recombinant AAV (rAAV) technology to rapidly develop a functional humanized mouse model, expressing the hACE2 specifically in the lung, i.e. the main entry organ of the SARS-CoV-2. We designed an expression cassette consisting of a cytomegalovirus (CMV) promoter driving the expression of human ACE2 and packaged it into rAAV vectors of serotype 9. Serotype 9 was chosen for its broad tropism, particularly its affinity for the lungs and central nervous system, as shown in Fig. 4. The vector was delivered into the lungs of wild-type (C57Bl/6NCrl) mice using a forced inhalation method, enabling localized expression of hACE2 and thereby humanizing the mice. The treated mice became susceptible to SARS-CoV-2 infection and exhibited COVID-19-like symptoms.

Our team's key contributions were the design, development, and validation of the AAV-CMV-hACE2 vector, followed by successful humanization of the mice and validation of the model. The following article outlines how this transient humanized mouse model was employed for the *in vivo* testing of bi- specific antibodies.

Article

Bispecific IgG neutralizes SARS-CoV-2 variants and prevents escape in mice

Raoul De Gasparo[1,16], Mattia Pedotti[1,16], Luca Simonelli[1], Petr Nickl[2], Frauke Muecksch[3], Irene Cassaniti[4], Elena Percivalle[4], Julio C. C. Lorenzi[5], Federica Mazzola[1], Davide Magri[6], Tereza Michalcikova[2], Jan Haviernik[7], Vaclav Honig[7,8], Blanka Mrazkova[2], Natalie Polakova[2], Andrea Fortova[7], Jolana Tureckova[2], Veronika Iatsiuk[2], Salvatore Di Girolamo[1], Martin Palus[7,8], Dagmar Zudova[2], Petr Bednar[7,9], Ivana Bukova[2], Filippo Bianchini[1], Dora Mehn[6], Radim Nencka[10], Petra Strakova[7], Oto Pavlis[11], Jan Rozman[2], Sabrina Gioria[6], Josè Camilla Sammartino[4], Federica Giardina[4], Stefano Gaiarsa[4], Qiang Pan-Hammarström[12], Christopher O. Barnes[13], Pamela J. Bjorkman[13], Luigi Calzolai[6], Antonio Piralla[4], Fausto Baldanti[4,14], Michel C. Nussenzweig[5,15], Paul D. Bieniasz[3,15], Theodora Hatziioannou[3], Jan Prochazka[2], Radislav Sedlacek[2], Davide F. Robbiani[1✉], Daniel Ruzek[7,8✉] & Luca Varani[1✉]

Neutralizing antibodies that target the receptor-binding domain (RBD) of the SARS-CoV-2 spike protein are among the most promising approaches against COVID-19[1,2]. A bispecific IgG1-like molecule (CoV-X2) has been developed on the basis of C121 and C135, two antibodies derived from donors who had recovered from COVID-19[3]. Here we show that CoV-X2 simultaneously binds two independent sites on the RBD and, unlike its parental antibodies, prevents detectable spike binding to the cellular receptor of the virus, angiotensin-converting enzyme 2 (ACE2). Furthermore, CoV-X2 neutralizes wild-type SARS-CoV-2 and its variants of concern, as well as escape mutants generated by the parental monoclonal antibodies. We also found that in a mouse model of SARS-CoV-2 infection with lung inflammation, CoV-X2 protects mice from disease and suppresses viral escape. Thus, the simultaneous targeting of non-overlapping RBD epitopes by IgG-like bispecific antibodies is feasible and effective, and combines the advantages of antibody cocktails with those of single-molecule approaches.

The COVID-19 pandemic has prompted substantial efforts to develop effective countermeasures against SARS-CoV-2. Preclinical data and phase-III clinical studies indicate that monoclonal antibodies could be effectively deployed for prevention or treatment during the viral symptoms phase of the disease[1,2]. Cocktails of two or more monoclonal antibodies are preferred over a single antibody as these cocktails result in increased efficacy and the prevention of viral escape. However, this approach requires increased manufacturing costs and volumes, which are problematic at a time when the supply chain is under pressure to meet the high demand for COVID-19 therapeutic agents, vaccines and biologics in general[4]. Cocktails also complicate formulation[5,6] and hinder strategies such as antibody delivery by viral vectors or by nonvectored nucleic acids[7,8]. One alternative is to use multispecific antibodies, which have the advantages of cocktails and single-molecule strategies.

To this end, we used structural information[9] and computational simulations to design bispecific antibodies that would simultaneously bind to (i) independent sites on the same RBD and (ii) distinct RBDs on a spike (S) trimer. We evaluated several designs using atomistic molecular dynamics simulations, and produced four constructs: of these, CoV-X2 was the most potent neutralizer of SARS-CoV-2 pseudovirus, and had a half-maximal inhibitory concentration (IC_{50}) of 0.04 nM (5.8 ng ml^{-1}) (Extended Data Fig. 1). CoV-X2 is a human-derived IgG1-like bispecific antibody in the CrossMAb format[10] that is the result of the combination of the Fragment antigen binding (Fab) of the monoclonal antibodies C121 and C135, which are two potent neutralizers of SARS-CoV-2[3]. Structural predictions showed that CoV-X2—but not its parental monoclonal antibodies—can bind bivalently to all RBD conformations on the S trimer, which prevents the binding of ACE2 receptor[11] (Fig. 1a, Extended Data Fig. 2).

CoV-X2 bound at a low nanomolar affinity to the RBD and S trimer of wild-type SARS-CoV-2 and to those of several naturally occurring SARS-CoV-2 variants, including B.1 (which contains D614G, in the S),

[1]Institute for Research in Biomedicine, Università della Svizzera italiana (USI), Bellinzona, Switzerland. [2]Czech Centre of Phenogenomics and Laboratory of Transgenic Models of Diseases, Institute of Molecular Genetics of the Czech Academy of Sciences, Vestec, Czech Republic. [3]Laboratory of Retrovirology, The Rockefeller University, New York, NY, USA. [4]Molecular Virology Unit, Microbiology and Virology Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy. [5]Laboratory of Molecular Immunology, The Rockefeller University, New York, NY, USA. [6]European Commission, Joint Research Centre, Ispra, Italy. [7]Veterinary Research Institute, Brno, Czech Republic. [8]Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, Ceske Budejovice, Czech Republic. [9]Faculty of Science, University of South Bohemia, Ceske Budejovice, Czech Republic. [10]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czech Republic. [11]Center of Biological Defense, Military Health Institute, Military Medical Agency, Techonin, Czech Republic. [12]Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden. [13]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [14]Department of Clinical Surgical Diagnostic and Pediatric Sciences, Università degli Studi di Pavia, Pavia, Italy. [15]Howard Hughes Medical Institute, The Rockefeller University, New York, NY, USA. [16]These authors contributed equally: Raoul De Gasparo, Mattia Pedotti. ✉e-mail: drobbiani@irb.usi.ch; ruzekd@paru.cas.cz; luca.varani@irb.usi.ch
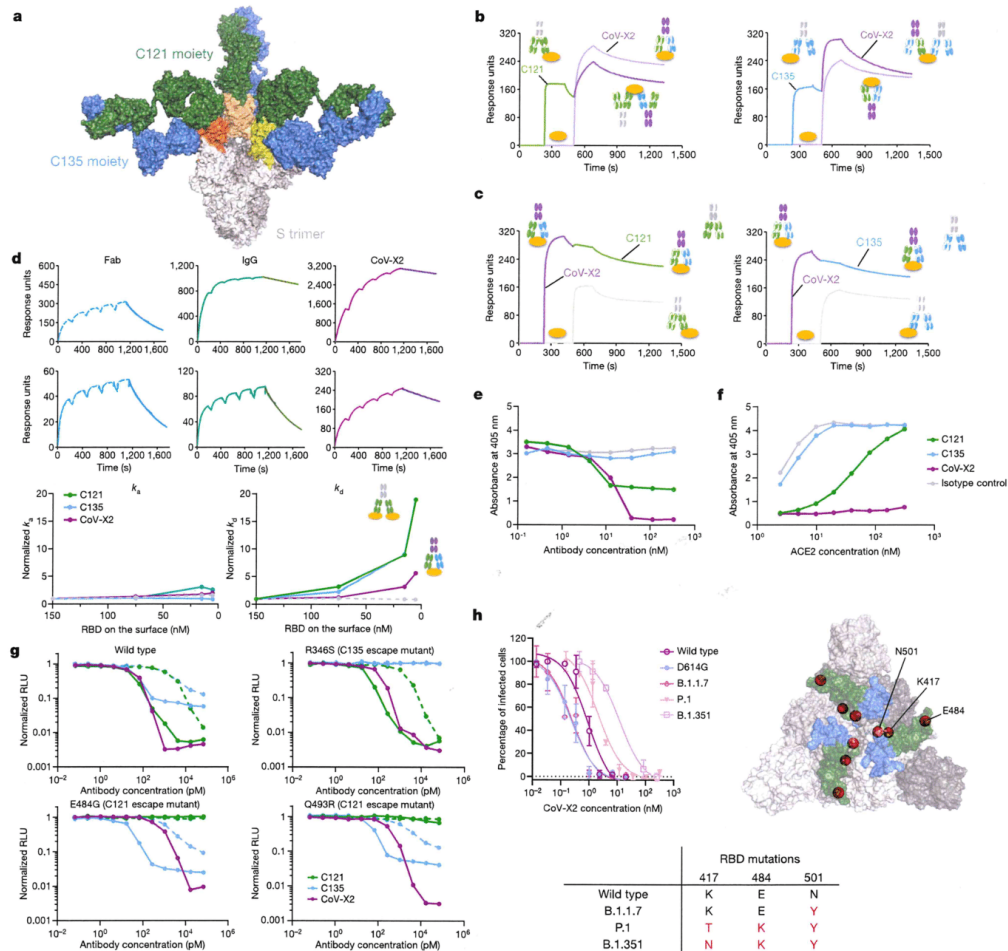
Fig. 1 | Biochemical and in vitro neutralizing properties of CoV-X2 are superior to those of its parental monoclonal antibodies. a, Computational simulations predict bivalent binding of CoV-X2 to all three RBDs on the S trimer (Extended Data Fig. 2). Green and blue, C121 and C135 moieties, respectively; yellow and orange, RBDs. **b**, **c**, Surface plasmon resonance (SPR) demonstrates that both arms of CoV-X2 are functional. In **b**, immobilized RBD complexed with C121 (left) or C135 (right) (first antibody) binds to CoV-X2 (second antibody). In **c**, RBD–CoV-X2 prevents binding by C121 (left) or C135 (right) single monoclonal antibodies. Paler colours denote controls (second antibody only). **d**, Both arms of CoV-X2 bind simultaneously to the RBD, as avidity is retained at decreasing RBD concentrations. Top and middle, representative SPR traces indicating the dissociations of Fab (left), IgG (centre) or CoV-X2 (right) binding to RBD immobilized at 150 nM (top) or 15 nM (middle) on the SPR chip (Extended Data Fig. 6). Bottom, plots of normalized k_a (left) and k_d (right) obtained with different concentrations of immobilized RBD. Increasing normalized k_d values indicate loss of avidity. Solid lines, IgG; dotted lines, Fab. **e**, **f**, CoV-X2 prevents ACE2 binding to S trimer in an ELISA. ACE2 binding to antibody–S-trimer complexes measured with increasing concentration of indicated antibody and constant ACE2 (**e**), or at constant antibody concentration with increasing ACE2 (**f**). Mean of two replicates shown. **g**, CoV-X2 neutralizes SARS-CoV-2 pseudovirus and escape mutants generated by its parental monoclonal antibodies. Normalized relative luminescence (RLU) for cell lysates after infection with NanoLuc-expressing SARS-CoV-2 pseudovirus in the presence of increasing concentrations of antibodies. Wild-type SARS-CoV-2 pseudovirus (top left) is shown alongside an escape mutant generated in the presence of C135 (R346S; top right), or two escape mutants generated in the presence of C121 (E484G; bottom left) and Q493R (bottom right)[14]. Dashed lines, parental Fabs; solid lines, IgG. Mean of two independent experiments with two replicates each. **h**, Neutralization of SARS-CoV-2 isolates with sequences corresponding to wild-type virus (which was first isolated in China), and the B.1 (D614G), B.1.1.7, P.1 and B.1.351 variants. Mean of three experiments with s.d. RBD residues mutated in the variants are indicated in the table and as red spheres on the S trimer structure, on which the epitopes of C135 (blue) and C121 (green) are shown.

B.1.1.7 (which contains N501Y, in the RBD) and B.1.351 (which contains K417N/E484K/N501Y, in the RBD)[12,13], as well as to escape mutants generated by the parental monoclonal antibodies[14] (Extended Data Figs. 3–5).

CoV-X2 also bound to preformed C121–RBD and C135–RBD complexes, which confirmed that both of its arms are functional (Fig. 1b, c). Next, we used an avidity assay by surface plasmon resonance to

43

Article

experimentally confirm the computational prediction that CoV-X2 can simultaneously engage two sites on the same RBD (Fig. 1d, Extended Data Fig. 6, Methods). Avidity occurs when IgG binds bivalently to antigens, which results in slower dissociation rates (k_d) (Extended Data Fig. 6a). Accordingly, C121 and C135 IgG showed avidity at high concentrations of antigen, owing to the intermolecular binding of adjacent RBDs; at lower concentrations of antigen, the k_d was faster as intermolecular binding was prevented by the increased distance between RBD molecules, which resulted in loss of avidity. Intramolecular avidity is not possible for C121 and C135 because a single epitope is available on each RBD molecule. By contrast, CoV-X2 maintained avidity even at low concentrations of antigen, which indicates that it undergoes bivalent, intramolecular binding (Fig. 1d, Extended Data Fig. 6). We then performed enzyme-linked immunosorbent assays (ELISAs) to evaluate the ability of CoV-X2 to inhibit the binding of recombinant ACE2 to the S trimer (Fig. 1e, f). Consistent with the structural information[9], C135 did not affect the interaction between ACE2 and S. C121, which occupies the ACE2-binding site on the RBD, prevented ACE2 binding only partially. By contrast, ACE2 binding was not detected in the presence of CoV-X2, which suggests a synergistic effect by the two moieties that comprise the bispecific antibody.

To assess the neutralizing ability of CoV-X2 in vitro, we first used SARS-CoV-2 pseudoviruses[15]. The bispecific antibody neutralized pseudovirus that carries wild-type SARS-CoV-2 S at sub-nanomolar concentrations ($IC_{50} = 0.04$ nM (5.8 ng ml^{-1}) and 90% inhibitory concentration of 0.3 nM (44 ng ml^{-1})), which was similar to or better than the parental IgG and an over 100-fold-better IC_{50} than that of the parental Fabs (Fig. 1g). CoV-X2 remained effective against pseudoviruses that bear escape mutations that make them resistant to the individual monoclonal antibodies[14] (Fig. 1g), and against a pseudovirus with RBD mutations that are found in the B.1.351 variant (which was first reported in South Africa) ($IC_{50} = 1.3$ nM (191 ng ml^{-1})) (Extended Data Fig. 5). To confirm the efficacy of CoV-X2, we performed plaque-reduction neutralization assays with infectious virus. CoV-X2 efficiently neutralized wild-type SARS-CoV-2 ($IC_{50} = 0.9$ nM), the B.1 variant that carries S(D614G), which was first detected in Europe ($IC_{50} = 0.2$ nM); the B.1.1.7 variant, which was first observed in the UK ($IC_{50} = 0.2$ nM); the P.1 variant, which was first isolated in Brazil ($IC_{50} = 2.1$ nM); and the B.1.351 variant ($IC_{50} = 12$ nM) (Fig. 1h). The P.1 and B.1.351 variants have almost identical mutations in the RBD; the only difference is the presence of an Asn (in B.1.351) versus Thr (in P.1) at position 417, which does not interact with CoV-X2. Nonetheless, the neutralization of B.1.351 was lower than that for P.1, which suggests that there are either some conformational differences in the RBD between the two variants or long-range effects that derive from additional mutations in the S. A similar behaviour is seen with the wild-type sequence (D614), which has a lower neutralization than S(D614G) even if no other difference is present; a plausible explanation is that a Gly at position 614 makes the CoV-X2 epitopes more accessible by favouring the 'up' conformation of the RBD[16]. We conclude that the in vitro binding and neutralizing properties of CoV-X2 make it preferable over its parental antibodies.

To assess the clinical potential of CoV-X2, we investigated its ability to protect against infection and disease in an animal model. We developed a mouse model in which human ACE2 (hACE2) is expressed by cells of the upper and lower respiratory tract after inhalation of a modified adeno-associated virus (AAV) (AAV-hACE2) (Fig. 2, Extended Data Fig. 7, Methods).

This approach enables the rapid production of large cohorts of mice, and has the advantage of being applicable to wild-type and mutant mouse colonies, independently of age and sex. Moreover, because AAV vectors are only weakly immunogenic and cytotoxic, the system allows for prolonged expression of hACE2[17–20] (Extended Data Fig. 7). SARS-CoV-2 infection of ACE2-humanized mice (hereafter, hACE2 mice) results in progressive weight loss, respiratory pathology and disease that require culling at 8 days post-infection (dpi) (Fig. 2a–c, Extended Data Fig. 7).

To evaluate the protective effect of antibodies, we treated hACE2 mice one day before SARS-CoV-2 infection (150 μg) (Fig. 2d, e, g, h) or 12 h after challenge (250 μg) (Fig. 2d, f) and monitored them over time. Upon intranasal infection with 1×10^4 plaque-forming units of SARS-CoV-2 (SARS-CoV-2/human/Czech Republic/951/2020), mice treated with an isotype control (an antibody against the envelope protein of Zika virus) showed weight loss starting at 3 dpi; by 8 dpi, most mice had lost approximately 25–30% of their body weight, reaching the end point for humane culling (Fig. 2e, f). We recovered infectious virus from the lungs (Fig. 2g), and detected viral RNA in the spleen (Fig. 2h) but not in the heart (data not shown). The lung pathology of these mice resembled that associated with severe COVID-19 in humans[21], and was characterized by diffuse alveolar damage (50–80% of tissue area), alveolar replacement with infiltrates of immune cells and fibroblasts, thickened septa and infiltrations by activated macrophages with foamy cytoplasm (Fig. 2j). By contrast, mice treated with CoV-X2 maintained their body weight ($P < 0.0001$ at 4–8 dpi, when compared to mice treated with isotype control) (Fig. 2e; the P values between all groups are given in Extended Data Table 1), had a reduced viral burden in lungs and viral RNA levels in spleen (Fig. 2g, h) and displayed neither macro- nor histopathological changes (diffuse alveolar damage of less than 5–10%) (Fig. 2i, j). Although infectious virus was detected at 2 dpi in all mice that were treated with C121 ($n = 5$), C135 ($n = 5$) or control ($n = 8$), in mice that received CoV-X2 it was detected only at low levels in one mouse (out of five) (Fig. 2g). At 5 dpi, we readily recovered infectious virus from control-treated mice (5 of 6 mice), but only from 1 out of 10 mice treated with CoV-X2 (Fig. 2g). Furthermore, CoV-X2 was also protective when administered at 12 h after SARS-CoV-2 challenge (Fig. 2d, f). As none of the mice treated with CoV-X2 exhibited symptoms at any time, we conclude that CoV-X2 protects mice from the disease.

Because monotherapy with C121 or C135 monoclonal antibodies leads to virus escape in vitro[14], we treated hACE2 mice with the individual antibodies and sequenced the virus. Only wild-type RBD sequences were obtained from control mice ($n = 10$). In mice treated with C121, there was selection for a mutation that results in the E484D substitution (5 of 5 mice that were analysed at 8 dpi). Escape mutations at E484 generated by C121 have previously been observed in vitro[14], and changes at this residue (which are also present in the B.1.351 and P.1 variants) reduce neutralization by human sera by more than tenfold[22]. The E484D substitution affects intermolecular hydrogen bonds at the core of the interface between C121 and the RBD, and it has previously been suggested to increase the affinity of the RBD for ACE2[23]. Virus containing Asp at position 484 is pathogenic, as 7 out of 9 mice treated with C121 developed disease (Fig. 2e) and only this virus was found in their lungs. By contrast (and unlike the in vitro results[14]), we did not observe virus evasion or pathology in mice treated with C135 ($n = 5$) (Fig. 2e and data not shown). Even though we did not retrieve infectious virus (at 8 dpi) from mice treated with CoV-X2 ($n = 13$) and noticed no symptoms in these mice, we detected low levels of residual viral RNA in some mice after 40 cycles of PCR amplification: in 6 of 13 mice, the virus sequence was wild type, and in 2 mice the overlapping sequencing traces were consistent with the coexistence of wild-type virus and a variant with Asp at position 484. Thus, in these 2 (out of 13) mice that contained the E484D variant, CoV-X2 remained protective even if the mutation diluted the effective antibody concentration (presumably leaving only the C135 moiety active).

Monoclonal antibodies that target the SARS-CoV-2 S are in advanced clinical trials and show promise against COVID-19[1,2]. Concomitant use of multiple antibodies is preferred for increased efficacy and adding resistance against viral evasion. Indeed, the virus can escape pressure by a single antibody in vitro and—as shown here—also in mice. Moreover, RBD mutations that threaten the efficacy of single monoclonal antibodies have already been detected in virus that is circulating in minks and humans, including mutations at the C121 and C135 epitopes (Extended Data Fig. 8). One disadvantage of antibody cocktails is the requirement
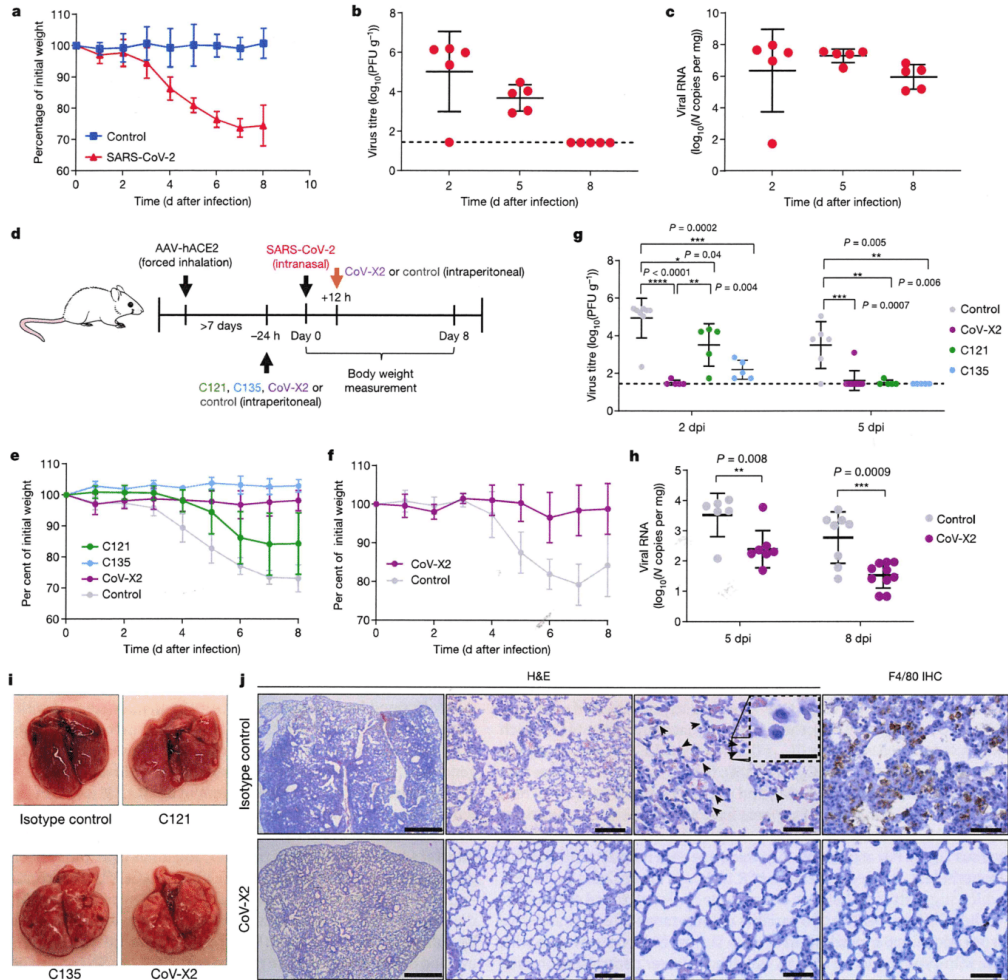
Fig. 2 | CoV-X2 protects mice transduced with AAV-hACE2 against SARS-CoV-2 disease. a, Body weight over time in mice infected with SARS-CoV-2. We transduced 13–15-week-old C57BL/6NCrl wild-type female mice with AAV-hACE2 by forced inhalation, delivering viral particles to the upper and lower respiratory tract. After >7 days, mice were infected with SARS-CoV-2 (1×10^4 plaque-forming units (PFU)) or received vehicle by the intranasal route. Weight was monitored daily for 8 days (SARS-CoV-2, $n = 5$; control, $n = 4$). Mean with s.d. **b**, Kinetics of viral burden in lungs of mice infected with SARS-CoV-2 ($n = 5$) by plaque assays. Mean with s.d. Dashed line, limit of detection. **c**, Kinetics of viral RNA levels in lung samples from mice infected with SARS-CoV-2 ($n = 5$) by quantitative PCR with reverse transcription (RT–qPCR). Mean with s.d. **d**, Wild-type mice were transduced with AAV-hACE2 by forced inhalation. After >7 days, mice were inoculated intraperitoneally with antibodies one day before (black arrow) or 12 h after (red arrow) being infected intranasally with SARS-CoV-2 (1×10^4 PFU). **e**, **f**, Body weight was monitored daily in mice treated 24 h before infection (**e**, C121, $n = 9$; C135, $n = 5$; CoV-X2,

$n = 13$; isotype control, $n = 10$) or 12 h after infection (**f**, CoV-X2, $n = 4$; control, $n = 10$). Mean with s.d. is shown. **g**, Lung viral burden by plaque assay at 2 dpi (isotype control, $n = 8$; CoV-X2, $n = 5$; C121, $n = 5$; C135, $n = 5$) and 5 dpi (isotype control, $n = 6$; CoV-X2, $n = 10$; C121, $n = 5$; C135, $n = 5$). The dashed line indicates the limit of detection; mean with s.d. P value was calculated with two-tailed Student's t-test. **h**, Viral RNA levels in the spleen by RT–qPCR at 5 and 8 dpi (isotype control, $n = 6$ or 8 (indicated by dots); CoV-X2, $n = 8$ or 10 (indicated by dots)). Mean with s.d. P value, two-tailed Student's t-test. **i**, Photographs of lungs from infected mice (8 dpi). **j**, Histopathology and F4/80 immunohistochemistry (IHC). Haematoxylin and eosin (H&E)-stained sections of paraffin-embedded lungs from infected mice (8 dpi). Arrowheads, foamy macrophages. F4/80 IHC shows abundant macrophage infiltration in lungs of mice treated with isotype control, but not those with CoV-X2. Each image representative of two separate experiments ($n = 3$–5 mice per group). Scale bars, 20 μm (inset in H&E), 50 μm (H&E right, F4/80 IHC), 100 μm (H&E middle), 1 mm (H&E, left).

45

Article

for a development and production capacity twice or more that required for single monoclonal antibodies, which is a considerable challenge in light of the increased demand owing to COVID-19 related vaccines and therapeutic agents (on top of the need to maintain production of biologics for other diseases)[4].

Multispecific antibodies offer the advantages of cocktails in a single molecule. We have shown that the CoV-X2 bispecific antibody is more effective than the related monoclonal antibodies at inhibiting ACE2 binding; it has a sub-nanomolar IC_{50} against a broader array of viral sequences and it protect mice from SARS-CoV-2 even when C121 (its potent parental monoclonal antibody) fails owing to viral escape. C135, the other parental monoclonal antibody, did not generate escape mutants in our mouse experiment, but readily generated them in vitro[14]. CoV-X2 is expected to be more resistant to viral escape as compared to monoclonal antibodies. We have shown that CoV-X2 binds and neutralizes mutants that are not recognized by its parental monoclonal antibodies, as well as variants of concern that have recently been detected in the UK[12], South Africa[13] and Brazil[24].

Unlike other multispecific antibodies[25], CoV-X2 is a fully human IgG-like molecule. As such, it is suitable for further development and could be engineered to alter effector functions. For example, the Fragment crystallizable (Fc) of CoV-X2 has already been modified by introducing LALA-PG mutations[26] to modulate its interaction with Fc receptor and complement without affecting its antigen-binding properties. The LALA modification prevents antibody-dependent enhancement of flavivirus infection[27] and may be a desirable modification in the context of SARS-CoV-2, as cellular and animal experiments with coronaviruses (including SARS-CoV[28,29]) support the possibility of antibody-dependent enhancement. Other modifications (such as the LS modification[26] for increased half-life) are easily achievable. Finally, CoV-X2 is human-derived and produced in the CrossMab format, which has already shown to be safe in clinical trials[30], which further supports the possibility of future development. IgG-like bispecific antibodies are therefore worth adding to the tools that we use to address SARS-CoV-2 and its future mutations.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03461-y.

1. DeFrancesco, L. COVID-19 antibodies on trial. *Nat. Biotechnol.* **38**, 1242–1252 (2020).
2. Klasse, P. J. & Moore, J. P. Antibodies to SARS-CoV-2 and their potential for therapeutic passive immunization. *eLife* **9**, e57877 (2020).
3. Robbiani, D. F. et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437–442 (2020).
4. Ecker, D. M. & Seymour, P. in *CPhI Annual Report 2020: Postulating the Post-COVID Pharma Paradigm*, 43–49 (Informamarkets, 2020).
5. Baum, A. et al. REGN-COV2 antibodies prevent and treat SARS-CoV-2 infection in rhesus macaques and hamsters. *Science* **370**, 1110–1115 (2020).
6. Schäfer, A. et al. Antibody potency, effector function, and combinations in protection and therapy for SARS-CoV-2 infection in vivo. *J. Exp. Med.* **218**, e20201993 (2021).
7. Schlake, T. et al. mRNA: a novel avenue to antibody therapy? *Mol. Ther.* **27**, 773–784 (2019).
8. Tiwari, P. M. et al. Engineered mRNA-expressed antibodies prevent respiratory syncytial virus infection. *Nat. Commun.* **9**, 3999 (2018).
9. Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).
10. Schaefer, W. et al. Immunoglobulin domain crossover as a generic approach for the production of bispecific IgG antibodies. *Proc. Natl Acad. Sci. USA* **108**, 11187–11192 (2011).
11. Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292.e6 (2020).
12. Kemp, S. et al. Recurrent emergence and transmission of a SARS-CoV-2 spike deletion ΔH69/V70. Preprint at https://doi.org/10.1101/2020.12.14.422555 (2020).
13. Tegally, H. et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. Preprint at https://doi.org/10.1101/2020.12.21.20248640 (2020).
14. Weisblum, Y. et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife* **9**, e61312 (2020).
15. Schmidt, F. et al. Measuring SARS-CoV-2 neutralizing antibody activity using pseudotyped and chimeric viruses. *J. Exp. Med.* **217**, e20201181 (2020).
16. Benton, D. J. et al. The effect of the D614G substitution on the structure of the spike glycoprotein of SARS-CoV-2. *Proc. Natl Acad. Sci. USA* **118**, e2022586118 (2021).
17. Han, K. et al. Lung expression of human ACE2 sensitizes the mouse to SARS-CoV-2 infection. *Am. J. Respir. Cell Mol. Biol.* **64**, 79–88 (2021).
18. Hassan, A. O. et al. A SARS-CoV-2 infection model in mice demonstrates protection by neutralizing antibodies. *Cell* **182**, 744–753.e4 (2020).
19. Sun, J. et al. Generation of a broadly useful model for COVID-19 pathogenesis, vaccination, and treatment. *Cell* **182**, 734–743.e5 (2020).
20. Sun, S. H. et al. A mouse model of SARS-CoV-2 infection and pathogenesis. *Cell Host Microbe* **28**, 124–133.e4 (2020).
21. Deshmukh, V., Motwani, R., Kumar, A., Kumari, C. & Raza, K. Histopathological observations in COVID-19: a systematic review. *J. Clin. Pathol.* **74**, 76–83 (2021).
22. Greaney, A. J. et al. Comprehensive mapping of mutations to the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human serum antibodies. *Cell Host Microbe* **29**, 463–476 (2021).
23. Chen, J., Wang, R., Wang, M. & Wei, G. W. Mutations strengthened SARS-CoV-2 infectivity. *J. Mol. Biol.* **432**, 5212–5226 (2020).
24. Sabino, E. C. et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* **397**, 452–455 (2021).
25. Dong, J. et al. Development of humanized tri-specific nanobodies with potent neutralization for SARS-CoV-2. *Sci. Rep.* **10**, 17806 (2020).
26. Saunders, K. O. Conceptual approaches to modulating antibody effector functions and circulation half-life. *Front. Immunol.* **10**, 1296 (2019).
27. Dejnirattisai, W. et al. Cross-reacting antibodies enhance dengue virus infection in humans. *Science* **328**, 745–748 (2010).
28. Yip, M. S. et al. Antibody-dependent infection of human macrophages by severe acute respiratory syndrome coronavirus. *Virol. J.* **11**, 82 (2014).
29. Yip, M. S. et al. Antibody-dependent enhancement of SARS coronavirus infection and its role in the pathogenesis of SARS. *Hong Kong Med. J.* **22** (Suppl 4), 25–31 (2016).
30. Klein, C. et al. Engineering therapeutic bispecific antibodies using CrossMab technology. *Methods* **154**, 21–31 (2019).

46

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

Protein expression and purification

To express SARS-CoV-2 S, a codon-optimized gene encoding residues 1–1208 (GenBank (https://www.ncbi.nlm.nih.gov/genbank/): MN908947) with proline substitutions at residues 986 and 987, a 'GSAS' substitution at the furin cleavage site (residues 682–685), a C-terminal T4 fibritin trimerization motif and a C-terminal 8× HisTag was synthesized and cloned into the mammalian expression vector pcDNA3.1(+) by Genscript. Codon-optimized nucleotide sequences encoding SARS-CoV-2 RBD (residues 319–591; GenBank: MN908947) with a C-terminal 8× HisTag was obtained from Genscript. hACE2 fused to the Fc region of mouse IgG at the C terminus was also synthesized and cloned into the mammalian expression vector pcDNA3.1(+). All proteins were produced by transient PEI transfection in Expi293F cells (ThermoFisher), purified from the cell supernatants 6 days after transfection by HiTrap Chelating HP (Cytiva) or HiTrap Protein A HP (Cytiva), analysed by SDS–PAGE and dynamic light scattering on a DynaPro NanoStar (Wyatt Technology, software Dynamics v.7.1.7.16). RBD and S trimer mutations were introduced in the plasmids by gene synthesis (Genscript) and purified. All proteins underwent quality control and biophysical characterization to ensure functionality, stability, lack of aggregation and batch-to-batch reproducibility.

Monoclonal antibody production. Plasmids containing the coding sequence for the production of the monoclonal antibodies C121, C135, and C144 were prepared as previously described[3]. These monoclonal antibodies were produced by transient PEI transfection in Expi293F cells (ThermoFisher), purified from the cell supernatants 6 days after transfection by HiTrap Protein A HP (Cytiva) and HiLoad Superdex 200 16/60 column (Cytiva).

Design, expression and purification of bispecific antibodies. Bispecific antibodies in the single-chain Fv format (CoV-scB1 and CoV-scB2) were designed from the sequences of the variable regions of monoclonal antibodies C144 and C135 (CoV-scB1) or C121 and C135[3] (CoV-scB2), following a previously described method[31], N-terminal signal peptides (residues 1–19; UniProt (https://www.uniprot.org/): P01743) and a C-terminal 6× HisTag were added. The constructs were synthesized and subcloned into the mammalian expression vector pcDNA3.1(+) by Genscript. The single-chain bispecific antibodies were produced by transient PEI transfection in Expi293F cells (ThermoFisher), purified from the cell supernatants 6 days after transfection by HiTrap Chelating HP (Cytiva) and HiLoad Superdex 75 16/60 column (Cytiva).

Bispecific antibodies in the CrossMAb format were designed from the sequences of the variable regions of antibodies C144 and C135 (CoV-X1) or C121 and C135[3] (CoV-X2). Light and heavy chain constant region sequences (UniProt P01834 and UniProt P01857) were added. The CrossMAbs were designed, as previously described[10], with CH1-CL crossover in the C135 moiety. Four constructs, one for each light chain (LC(C144), LC(C121) and LC^{CH1-CL}(C135)) and heavy chain (HC(C144), HC(C121) and HC^{CH1-CL}(C135)) were synthesized and subcloned into the mammalian expression vector pcDNA3.1(+) by Genscript. Signal peptides were included at the N terminus of the variable sequences for expression purposes (residues 1–19; UniProt P01743 for the heavy chains and residues 1–20; UniProt P06312 for the light chains). The antibodies were produced by PEI transient transfection, plasmid ratio 1:1:1:1, in Expi293F cells (ThermoFisher), purified from the cell supernatants 6 days after transfection by HiTrap Protein A HP (Cytiva), and HiLoad Superdex 200 16/60 column (Cytiva).

All antibodies underwent quality control and biophysical characterization to ensure functionality, stability, lack of aggregation and batch-to-batch reproducibility.

Computational modelling

CoV-scB1 and CoV-scB2 were modelled starting from the variable fragment of available experimental structures of the parental antibodies (Protein Data Bank (PDB) (https://www.rcsb.org/) identifiers: 7K8X for C121, 7K8Z for C135 and 7K90 for C144); the connecting linkers were manually added (Pymol); the stability and feasibility of the bispecific constructs on the S trimer were manually and visually investigated according to structural biology considerations.

CoV-X1 and CoV-X2 were assembled starting from the experimental structures of individual Fab antibodies in complex with SARS-CoV-2 (PDB 7K8X for C121, 7K8Z for C135 and 7K90 for C144). The Fc moiety was manually placed in proximity to the C terminus of each Fab heavy chain; CH1 domains were then connected to the Fc using the ALMOST toolkit[32], thus obtaining the full antibody structures.

The S trimer coordinates were obtained from available experimental structures; loop regions not present in the structure were modelled using the I-TASSER suite[33]. PDB 6VXX served as a basis for the 'three down' conformation; PDB 6VYB for the 'one RBD up' conformation; and PDB 7A93 for the 'two RBD up' conformation. Conformations not directly available (for example, trimer with 'three RBD up') were obtained by structural alignment and repetition of the appropriate conformation in the S monomer structures (for example, monomeric 'RBD up') using the PyMol software suite[34]

All the combinations of S trimer (three RBD down, one up, two up and three up) and antibodies (bispecific antibodies formed by combinations of C121, C135 and C144) were subjected to 400 ns or 200 ns of fully atomistic molecular dynamics simulations to obtain energetically favourable and stable conformations using GROMACS[35]. Calculations were performed on the CINECA-Marconi100 supercomputer.

The system was initially set up and equilibrated through standard molecular dynamics protocols: proteins were centred in a triclinic box, 0.2 nm from the edge, filled with SPCE water model and 0.15 M Na^+Cl^- ions using the AMBER99SB-ILDN protein force field. Energy minimization was performed to let the ions achieve a stable conformation. Temperature and pressure equilibration steps, respectively, at 310 K and 1 bar of 100 ps each were completed before performing the full molecular dynamics simulations with the above-mentioned force field. Molecular dynamics trajectory files were analysed after removal of periodic boundary conditions. The stability of each simulated complex was verified by root mean square deviation and visual analysis.

SPR

The antibody binding properties were analysed at 25 °C on a Biacore 8K instrument (GE Healthcare) using 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.005% Tween-20 as running buffer.

SARS-CoV-2 RBD or full S, and their mutants, were immobilized on the surface of CM5 chips (Cytiva) through standard amine coupling. Increasing concentrations of antibody (3.12, 6.25, 12.5, 25 and 50 nM) were injected using a single-cycle kinetics setting; analyte responses were corrected for unspecific binding and buffer responses. Curve fitting and data analysis were performed with Biacore Insight Evaluation Software v.2.0.15.12933.

Competition experiments were performed to assess the ability of COV-X2 to bind its target with both arms. A low amount of RBD (5 nM) was immobilized on the surface of a CM5 chip through standard amine coupling. C121 or C135 antibodies were injected at a high concentration (1.5 µM) to saturate the corresponding binding sites on the RBD; CoV-X2 (200 nM) was subsequently injected. The same experimental setting was performed with a different injection order as a control: COV-X2 (1.5 µM) injected to saturate RBD binding sites and subsequent injection of C121 or C135 (200 nM).

Article

Analysis and comparison of kinetics parameters at different RBD concentrations were performed as following. RBD was immobilized on the surface of a CM5 chip at 5, 15, 75 and 150 nM. Increasing concentrations of antibody (3.12, 6.25, 12.5, 25 and 50 nM) were injected using a single-cycle kinetics setting; analyte responses were corrected for unspecific binding and buffer responses. Curve fitting and data analysis were performed with Biacore Insight Evaluation Software v.2.0.15.12933.

Binding inhibition of hACE2

ELISAs were used to investigate the ability of antibodies to inhibit the binding of S to hACE2. Each experiment was performed in duplicate, reporting the mean of the two replicates; error bars represent the s.d. of the measured values. We used 96-well ELISA plates, coated at 4 °C with 37 nM S, and washed and blocked with PBS + 10% FCS. Antibodies were then added either at constant saturating concentration (160 nM) or at different dilutions (starting from 340 nM and serially diluted 1 to 3) and incubated 1 h at 25 °C; after washing, hACE2–mouse Fc was added either at constant saturating concentration (160 nM) or at different dilutions (starting from 340 nM and serially diluted 1 to 2) and left 1 h at 25 °C. After further washing, bound hACE2 was detected using standard protocols with goat anti-mouse IgG coupled to alkaline phosphatase (dilution 1:500, SouthernBiotech). ELISA plates were measured with the reader software Gen5 version 1.11.5 (BioTek Instruments). Data were analysed with Microsoft Excel 2016 and GraphPad Prism version 8.4.2.

SARS-CoV-2 pseudotyped reporter virus and pseudotyped-virus neutralization assay

A panel of plasmids expressing RBD-mutant SARS-CoV-2 S in the context of pSARS-CoV-2-S$_{\Delta19}$ have previously been described[14,36]. The mutant 'KEN' (K417N/E484K/N501Y) was constructed in the context of a pSARS-CoV-2-S$_{\Delta19}$ variant with a mutation in the furin cleavage site (R683G). The IC$_{50}$ of these pseudotypes were compared to a wild-type SARS-CoV-2 S sequence carrying R683G. Generation of SARS-CoV-2 pseudotyped HIV-1 particles was performed as previously described[3,15].

The neutralization activity of the bispecific and monoclonal antibodies was measured as previously reported[3,15]. In brief, fourfold serially diluted antibodies were incubated with SARS-CoV-2 pseudotyped virus for 1 h at 37 °C. The mixture was subsequently incubated with 293T cells expressing ACE2 for 48 h; the cells were washed twice with PBS and lysed with Luciferase Cell Culture Lysis 5× reagent (Promega). NanoLuc activity in lysates was measured using the Nano-Glo Luciferase Assay System (Promega) with the Modulus II Microplate Reader User interface (TURNER BioSystems). The obtained relative luminescence units (which are indicative of virus presence) were normalized to those derived from cells infected with SARS-CoV-2 pseudotyped virus in the absence of antibodies. The IC$_{50}$ was determined using four-parameter nonlinear regression (GraphPad Prism).

SARS-CoV-2 virus-neutralization assay

The neutralizing activity of CoV-X2 against SARS-CoV-2 wild type, B.1 (D614G), B.1.1.7, P.1 and B.1.351 was investigated by plaque-reduction neutralization tests following a previously reported protocol[37]. In brief, 50 µl of antibody, starting from a concentration of 12 µg ml^{-1} or 190 µg ml^{-1} in a serial threefold dilution, were mixed in a flat-bottomed tissue culture microtitre plate (COSTAR) with an equal volume of 100 median tissue culture infectious dose of infectious virus that was isolated from patients with COVID-19, sequenced, titrated and incubated at 33 °C in 5% CO$_2$. After 1 h, 3×10^4 (100 µl) Vero E6 cells (VERO C1008, Vero 76, clone 18 E6, Vero E6; ATCC CRL-1586) were added to each well. After 3 days of incubation, cells were stained with Gram's crystal violet solution (Merck) plus 5% formaldehyde 40% m/v (Carlo Erba S.p.A.) for 30 min. Microtitre plates were then washed in water. Wells were analysed to evaluate the degree of cytopathic effect compared to untreated control. Each experiment was performed in triplicate.

The IC$_{50}$ was determined using three-parameter nonlinear regression (GraphPad Prism).

AAV-hACE2 vector design

Plasmid design and construction. The AAV transfer plasmid expressing hACE2 (AAV-hACE2) was created by replacing the GFP sequence with *ACE2* cDNA obtained from the HEK293 cell line (no. CRL-1573, ATCC, mycoplasma-free, population doubling lower than 13) in a pAAV-GFP plasmid (no. AAV-400, Cellbiolabs). HEK293 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (no. D5796, Sigma Aldrich) supplemented with 10% FBS (no. 10082139, Gibco) and 1% penicillin–streptomycin (no. XC-A4122100, BioSera). Cells were kept at 37 °C under 5% CO$_2$ atmosphere and 95% humidity.

To obtain the hACE2 sequence, total RNA was isolated from a confluent 10-cm^2 plate of HEK293 cells using RNeasy Mini Kit (Qiagen, no. 74104) according to the manufacturer's protocol and reverse-transcribed with M-MLV Reverse Transcriptase (Promega, no. M1701). The generated cDNA was used as a template for PCR amplification with a pair of primers (*ACE2* forward: 5′ATGTCAAGCTCTTCCTGG 3′, *ACE2* reverse: 5′CTAAAAGGA GGTCTGAACATC 3′) specific for hACE2 (NM_021804.3) using Phusion High-Fidelity DNA Polymerase (NEB, no. M0530S). The PCR product was separated in a 1% agarose gel (SeaKem LE AGAROSE; East Port, no. 50004); the band of appropriate size (2,418 bp) was extracted using NucleoSpin Gel and PCR Clean-up (Takara, no. 740609) according to the manufacturer protocol. The extracted product was treated with DreamTaq Green DNA Polymerase (Thermo Scientific, no. EP0711) in the presence of dATP to add 3′A overhangs to the PCR product. The product was then subcloned into pGEM-T Easy Vector (Promega, no. A1360). Proper insertion of the product was assessed by HindIII (Thermo Scientific, no. ER0501) and SacI (Thermo Scientific, no. ER1132) double-digestion control. Kozak sequence and SpeI recognition site were added at the 5′ and 3′ end of the *ACE2* cDNA PCR product, respectively, by amplification with a specific pair of primers (hACE2_Kozak_Fw: 5′-CAGGGGACGATGTCAAGCTCTTCCTGG -3′, hACE2_Spe_Rv: 5′-ACTAGTGATCTAAAAGGAGGT-3′) using Phusion High-Fidelity DNA Polymerase (NEB, no. M0530S). The amplified product was separated in 1% agarose gel (SeaKem LE AGAROSE; East Port, no. 50004), extracted with NucleoSpin Gel and PCR Clean-up (Takara, no. 740609) according to the manufacturer's protocol. The extracted *ACE2* sequence was PCR-amplified with a pair of primers with microhomology arms (hACE2_IF_Fw: 5′-TTCGAACATCGATTGCAGGGGACGATGTCAAG-3′, hACE2_Spe_IF_Rv: 5′-GCGCTGCTCGAGGCAACTAGTGATCTAAAAGGAGGT -3′) and Phusion High-Fidelity DNA Polymerase (NEB, no. M0530S), and was subsequently purified from agarose gel with NucleoSpin Gel and PCR Clean-up (Takara, no. 740609).

The GFP sequence in the pAAV-GFP control vector (no. AAV-400, Cellbiolabs) was excised by double digestion using EcoRI (Thermo Scientific, no. ER0271)/HindIII (Thermo Scientific, no. ER0501) restriction enzymes and replaced with the hACE2 sequence flanked by microhomology arms using the In-Fusion cloning system (In-Fusion HD Cloning Kit, Takara Bio Europe, no. 639648) according to the manufacturer's protocol. Proper insertion and presence of the SpeI recognition site were confirmed by double digestion using HindIII (Thermo Scientific, no. ER0501)/MluI (Thermo Scientific, no. ER0562) and SpeI (Thermo Scientific, no. ER1252), respectively. The generated AAV-hACE2 vector was sequenced (Eurofins Genomics) using the following primers: CMV_Fw: 5′-AAATGGGCGGTAGGCGTG-3′, seq_hACE2_start1: 5′-TGGAGATCTGAGGTCGG-3′, seq_hACE2_start2: 5′-TCTTCCTCCCACAGCTCCT-3′, seq_hACE2_1: 5′-CAGTTGATT GAAGATGTGGA-3′, seq_hACE2_2: 5′-AGAAGTGGAGGTGGATG-3′, seq_hACE2_3: 5′-AGAACTGAAGTTGAAAAGG-3′. The produced hACE2 sequence was 100% identical to the reference hACE2 sequence (NM_021804.3).

Transfection of Neuro-2a cells. Neuro-2a cells (CCL-131, ATCC, mycoplasma-free, population doubling lower than 10) used for

validation of AAV-hACE2 function were cultured in DMEM (no. D5796, Sigma Aldrich) supplemented with 10% FBS (no. 10082139, Gibco) and 1% penicillin–streptomycin (no. XC-A4122100, BioSera) and kept at 37 °C under 5% CO_2 atmosphere and 95% humidity. Three 60-mm^2 plates were seeded each with 4×10^5 cells and transfected the next day with the AAV-hACE2 vector using Lipofectamine 2000 Transfection Reagent (Invitrogen, no. 11668027) according to the manufacturer's protocol for the 6-well plate transfection. Transfection with pAAV-GFP control vector (no. AAV-400, Cellbiolabs) and non-transfected HEK293 cells were used as negative controls. After 48 h, cells were collected for western blotting in RIPA buffer, supplemented with AEBSF protease inhibitor (AppliChem, no. A1421) and cOmplete Mini Protease Inhibitor Cocktail (Sigma Aldrich, no. 04693124001). Expression of hACE2 in Neuro-2a cells was compared to the non-transfected HEK293 cells.

AAV-hACE2 particle production

AAV293 cell transfection. The AAV293 cell line (Cellbiolabs, no. AAV-100, mycoplasma-free, population doubling lower than 8) used for AAV production was cultured in DMEM (no. D5796, Sigma Aldrich) supplemented with 10% FBS (no. 10082139, Gibco), 1% penicillin–streptomycin (no. XC-A4122100, BioSera) and 1% NEAA (no. M7145, Sigma Aldrich). Cells were kept at 37 °C under 5% CO_2 atmosphere and 95% humidity. A day before transfection, 6.5×10^6 cells were seeded on 15-cm^2 cultivation plates to reach 80–90% confluency on the day of transfection. Vectors pHelper (no. 340202, Cell BioLabs), AAV Rep/Cap 2/9n (no. 112865, Addgene) and AAV-hACE2 were used for transfection in equimolar ratio. The total amount of DNA (28 µg per plate) diluted in 1 ml per plate of DMEM was mixed with linear polyethylenimine hydrochloride, M.W. 40000 PEI (no. 24765-1, Polysciences) in 1:2.7 ratio. After 20 min of incubation at room temperature, the transfection mixture was added to a cultivation plate with FBS-reduced medium (DMEM supplemented with 1% FBS) in a dropwise manner. After 5 h of incubation at 37 °C under 5% CO_2 atmosphere, the medium was removed and replaced with fresh complete growth medium (DMEM supplemented with 10% FBS and 1% penicillin–streptomycin).

AAV293–hACE2 collection. Three days after transfection, both cell medium and cells were collected for AAV particle isolation. These procedures were adapted and modified from a previous publication[38]. The medium was collected into 50-ml centrifuge tubes and cells were washed twice with 5 ml of PBS. Subsequently, cells were scraped in 1 ml of PBS and centrifuged at 1,000g for 10 min at 4 °C. The supernatant was added to the previously collected cell medium; the cell pellet was kept on ice during subsequent processing. The medium was centrifuged at 3,200g for 15 min at 4 °C and the supernatant was then filtered into a sterile glass bottle using 0.22-µm PES membranes. PEG-8000 (no. V3011, Promega) was added to the medium in a glass bottle in a 1:4 ratio. The mixture was stirred slowly at 4 °C for 1 h and then incubated overnight at 4 °C without stirring to allow full virus precipitation. The following day, the medium was centrifuged at 2,800g for 20 min at 4 °C and the pellet resuspended in 10 ml of PBS solution with 0.001% Pluronic F-68 non-ionic surfactant (no. 24040032, Gibco) and 200 mM NaCl (no. S5886, Sigma Aldrich) and sonicated at 50% amplitude with 4× 1-s on/15-min off pulses on ice. The cell lysate was centrifuged at 3,200g for 15 min at 4 °C. Subsequently, 50 U ml^{-1} of benzonase nuclease (no. E1014-25KU, Sigma Aldrich) was added to the viral suspension to degrade any residual DNA. After incubation for 1 h at 37 °C, the viral suspension was centrifuged at 2,400g for 10 min at 4 °C and the clarified supernatant was further purified.

AAV-hACE2 purification by iodixanol gradient ultracentrifugation. A gradient consisting of 15% iodixanol (in 1 M NaCl, 2.7 mM $MgCl_2$ 2 mM KCl in phosphate buffer), 25% iodixanol (in 2.7 mM $MgCl_2$, 2 mM KCl, and 0.001% phenol red in PBS), 40% iodixanol (in 2.7 mM $MgCl_2$ and 2 mM KCl in PBS) and 0.002% phenol red (no. P3532, Sigma Aldrich) in

60% iodixanol (OptiPrep Density Gradient Medium, no. D1556, Sigma Aldrich) was prepared in QuickSeal tubes according a previous publication[38]. Five ml of clarified viral supernatant was carefully added on the top of the gradient and the rest of the tube was filled up with PBS. Ultracentrifugation was carried out at 350,000g for 90 min in a pre-cooled T70i rotor at 10 °C. After ultracentrifugation, approximately 750-µl fractions were collected from the 40% iodixanol phase using an 18G needle puncturing the QuickSeal tube at the interface of the 60% and 40% iodixanol.

AAV-hACE2 purity validation and buffer exchange. The purity of the collected fractions from 40% iodixanol containing AAV-hACE2 particles was assessed by SDS–PAGE. Ten µl of each collected fraction was mixed with 3.5 µl of 4× Laemmli buffer and loaded to 4–20% Mini-PROTEAN TGX Precast Protein Gel (no. 4561096, Bio-Rad). The gels were briefly washed in dH$_2$O and stained with silver according to the manufacturer's protocol (Pierce Silver Stain Kit, no. 10096113, Thermo Scientific).

Selected AAV-hACE2 fractions were pooled and concentrated using Amicon Ultra-0.5 Centrifugal Filter Unit (molecular weight cut-off of 100 kDa). First, Amicon filter membranes were activated by incubation with 0.1% pluronic F-68, 0.01% pluronic F-68 and 200 mM NaCl, followed by 0.001% of pluronic F-68 in PBS and centrifugation at 1,900g for 5 min at 4 °C. Pooled fractions with AAV-hACE2 particles were loaded onto activated Amicon filter membranes and centrifuged at 2,600g for 5 min at 4 °C. The membranes were washed several times with 0.001% pluronic F-68 in PBS (centrifugation at 2,600g for 8 min at 4 °C) until the residual iodixanol was completely removed from the solution. To elute and concentrate the viral suspension, the membranes were covered with about 5 ml of formulation buffer and incubated for 5 min at room temperature. Amicon filters were centrifuged at 2,600g at 4 °C for approximately 1.5 min until around 0.5 ml of the formulation buffer with AAV-hACE2 particles was left. The eluate was transferred into sterile 1.5-ml tubes, quantified and stored at 4 °C for up to 2 weeks for short-term in vivo application or at −80 °C for long-term storage.

AAV-hACE2 titration by qPCR. The protocol for quantification and determination of the number of genome-containing particles of AAV-hACE2 was adapted from a previous publication[39], using qPCR. Purified AAV-hACE2 particles were treated with DNase I (no. EN0521, Thermo Scientific) to eliminate contaminating plasmid DNA. Serial dilutions of a AAV-hACE2 viral suspension were used as template in two separate reactions, one detecting viral inverted terminal repeat (ITR) sequences (ITR_Fw: 5′-GGAACCCCTAGTGATGGAGTT-3′, ITR_Rv: 5′-CGGCCTCAGTGAGCGA-3′) and the second hACE2 (hACE2_Fw: 5′-CCATTGGTCTTCTGTCACCCG-3′, hACE2_Rv: 5′-AGACCATCCACCTCCACTTCTC-3′). Data analysis was performed using the LightCycler 480 Software, version 1.5. AAV concentration (the number of viral genomes in 1 µl of AAV sample) was determined by comparison to standard curves of defined concentrations of AAV-hACE2 vector. Each qPCR run was performed in triplicate; six serial dilutions of the AAV-hACE2 vector were used as positive controls and standards

Mouse experiments

This study was carried out in strict accordance with the Czech national laws and guidelines on the use of experimental animals and protection of animals against cruelty (Animal Welfare Act No. 246/1992 Coll.). The protocol was approved by the Committee on the Ethics of Animal Experiments of the Institute of Parasitology, Institute of Molecular Genetics of the Czech Academy of Sciences, and of the Departmental Expert Committee for the Approval of Projects of Experiments on Animals of the Academy of Sciences of the Czech Republic (permits 82/2020 and 101/2020).

Application of AAV-hACE2 viral particles to mice. Thirteen-to-fifteen-week-old C57BL/6NCrl female mice were anaesthetized by

Article

intraperitoneal injection of ketamine and xylazine (0.1 mg per g body weight) (Biopharm) and 0.01 mg per g (Bioveta), respectively). Viral particles containing AAV-hACE2 were diluted to a final concentration of 4×10^9 genome copies in 40 μl of PBS. This volume was applied to mice by forced inhalation. The tip of the nose was gently clipped with tweezers and the tongue gently pulled out. After the mouse started breathing through the oral cavity, 40 μl of viral suspension were applied by a 200-μl pipette tip into the oral cavity and inhaled by the mouse through the trachea into the lungs.

To access whether hACE2 was expressed in lungs, lung tissue was collected and analysed by western blot in RIPA buffer, supplemented with AEBSF protease inhibitor (AppliChem, no. A1421) and cOmplete Mini Protease Inhibitor Cocktail (Sigma Aldrich, no. 04693124001) 1, 2 and 4 weeks after application. Expression of hACE2 in AAV-hACE2 transduced mice was compared to nontreated C57BL/6NCrl mice. Histone H3 antibody (cat. no.: ab1791, Abcam, lot: GR3237685-2; 1:1,000 dilution) was used as a loading control.

Mouse infection. SARS-CoV-2 (strain SARS-CoV-2/human/Czech Republic/951/2020, isolated from a clinical sample at the National Institute of Public Health, Prague), provided by J. Weber, was used for mouse infection. The virus was passaged in Vero E6 cells five times before its use in this study.

At least 7 days after application of the AAV-hACE2 virus particles, mice were infected intranasally with SARS-CoV-2 (1×10^4 plaque-forming units) in a total volume of 50 μl DMEM. Mice were monitored and weighted daily over an eight-day period. Treated mice were injected intraperitoneally with either 150 μg of antibodies 24 h before the infection or 250 μg 12 h after infection. Mice were killed at the indicated times after infection and their tissues collected for analysis.

Measurement of the viral burden. Tissues were weighed and homogenized using Precellys 24 (Bertin Technologies) and prepared as 20% (w/v) suspension in DMEM containing 10% newborn calf serum. The homogenates were clarified by centrifugation at 5,000g and the supernatant medium was used for plaque assay and viral RNA isolation.

Plaque assays were performed in Vero E6 cells (ATCC CRL-1586; mycoplasma-free) grown at 37 °C and 5% CO_2 in DMEM (no. LM-D1112/500, Biosera) supplemented with 10% FBS (no. FB-1001G/500; Biosera), and 100 U ml^{-1} penicillin, 100 μg ml^{-1} streptomycin (Antibiotic Antimycotic Solution; no. A5955; Sigma), and 1% L-glutamine (no. XC-T1755/100; Biosera) using a modified version of a previously published protocol[40]. In brief, serial dilutions of virus were prepared in 24-well tissue culture plates and cells were added to each well (0.6×10^5–1.5×10^5 cells per well). After 4 h, the suspension was overlaid with 1.5% (w/v) carboxymethylcellulose (no. C4888; Sigma) in DMEM. Following a 5-day incubation at 37 °C and 5% CO_2, plates were washed with phosphate-buffered saline and the cell monolayers were stained with naphthol blue black (no. 195243; Sigma). The virus titre was expressed as plaque-forming units per ml.

RNA was isolated from tissue homogenates using the QIAmp Viral RNA mini kit (no. 52906; Qiagen) following manufacturer's instructions.

Viral RNA was quantified using EliGene COVID19 Basic a RT (no. 90077-RT-A; Elisabeth Pharmacon) according to the manufacturer's protocol. A calibration curve was constructed from four dilutions of a sample that was quantified using Quanty COVID-19 kit (no. RT-25; Clonit), according to the recommendations from the manufacturer. All real-time PCR reactions were performed using a LightCycler 480 (Roche).

For sequencing, isolated RNA was used as a template for one-step RT-PCR (Qiagen OneStep RT-PCR Kit; no. 210212; Qiagen) with primers specific for SARS-CoV-2 RBD sequence (SARS-CoV-2_seq_FW: 5′-GCACTTGACCCTCTCTCAGAAAC-3′; SARS-CoV-2_seq_RV: 5′-GACTCAGTAAGAACACCTGTGCC-3′). The reaction mixture (final volume 25 μl) contained 5 μl of QIAGEN OneStep RT-PCR buffer, 1 μl of dNTP mix, 5 μl of 5× Q-solution, 1 μl of QIAGEN OneStep RT-PCR enzyme mix, 6 μl of RNase-free water, 1.5 μl of each primer (stock concentration, 0.01 mM) and 4 μl of template RNA. The cycling conditions were as follows: reverse transcription (30 min at 50 °C), initial PCR activation (15 min at 95 °C), 3-step cycling: 40 cycles of 94 °C for 30 s, 52.6 °C for 30 s, and 72 °C for 1 min, followed by final extension (10 min at 72 °C). The PCR products were visualized in a 1.7% agarose gel in Tris-acetate-EDTA buffer. The amplified DNA was purified by using Wizard SV Gel and PCR Clean-Up System (no. A9285; Promega), according to the recommendations of the manufacturer.

The purified DNA was prepared for sequencing (Sanger method) by a commercial service (Eurofins Genomics) with the following conditions (final volume 17 μl): 15 μl of PCR product with a concentration of 5 ng μl^{-1} and 2 μl of primer with a concentration of 10 μM. The sequencing data were analysed using BioEdit Sequence Alignment Editor, version 7.2.0.

Histology and immunohistochemistry. Lungs were fixed in 4% PFA. Tissues were processed using a Leica ASP6025 automatic vacuum tissue processor and embedded in paraffin using a Leica EG1150 H+C embedding station. Two-μm sections were prepared using a Leica RM2255 rotary microtome and sections were stained with H&E using Leica ST5020 + CV5030 stainer and coverslipper.

To assess the presence of macrophages, a rabbit anti-mouse F4/80 monoclonal antibody (D2S9R XP^R rabbit monoclonal antibody, cat. no. 70076, Cell Signaling Technology, lot 5, RRID AB_2799771) was used at 1:800 dilution as primary antibody. The histological sections (thickness 4–5 μm) were deparaffinized in a Multistainer Leica ST5020 (Leica Biosystems). Antigens were retrieved by heating the slides in citrate buffer pH 6 (Zytomed Systems). Endogenous peroxidase was neutralized with 3% H_2O_2. Sections were incubated for 1 h at room temperature with a 1:800 dilution of the primary antibody. After washing they were incubated with anti-rabbit secondary antibody conjugated with HRP (Zytomed, cat. no. ZUC 032-100, lot A0880-4; no dilution). Staining of the sections was developed with a diaminobenzidine substrate kit (DAKO, Agilent) and sections were counterstained with Harris haematoxylin (Sigma Aldrich, Merck) in a Multistainer Leica.

Biosafety statement

All work with infectious SARS-CoV-2 was performed in biosafety level 3 facilities at the Institute of Parasitology, Biology Centre of the Czech Academy of Science (Ceske Budejovice) and Veterinary Research Institute (Brno) using appropriate powered air-purifying, positive pressure respirators and protective equipment.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data that support the findings of this study are available within the Article and its Supplementary Information. Any other data are available from the corresponding author upon reasonable request. Published data were taken from GenBank (https://www.ncbi.nlm.nih.gov/genbank/), UniProt (https://www.uniprot.org/), PDB (https://www.rcsb.org/) and the ViPR database (https://www.viprbrc.org/). Source data are provided with this paper.

31. Bardelli, M. et al. A bispecific immunotweezer prevents soluble PrP oligomers and abolishes prion toxicity. *PLoS Pathog.* **14**, e1007335 (2018).
32. Fu, B. et al. ALMOST: an all atom molecular simulation toolkit for protein structure determination. *J. Comput. Chem.* **35**, 1101–1105 (2014).
33. Yang, J. et al. The I-TASSER suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
34. Schrodinger. The PyMOL Molecular Graphics System, Version 1.8 (Schrodinger 2015).
35. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

36. Wang, Z. et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* (2021).
37. Percivalle, E. et al. West Nile or Usutu virus? A three-year follow-up of humoral and cellular response in a group of asymptomatic blood donors. *Viruses* **12**, 157 (2020).
38. Zolotukhin, S. et al. Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther.* **6**, 973–985 (1999).
39. Aurnhammer, C. et al. Universal real-time PCR for the detection and quantification of adeno-associated virus serotype 2-derived inverted terminal repeat sequences. *Hum. Gene Ther. Methods* **23**, 18–28 (2012).
40. De Madrid, A. T. & Porterfield, J. S. A simple micro-culture method for the study of group B arboviruses. *Bull. World Health Organ.* **40**, 113–121 (1969).

a

CoV-scB1

CoV-scB2

CoV-X1

CoV-X2

b

Extended Data Fig. 1 | Neutralization of SARS-CoV-2 pseudovirus by bispecific antibodies. a, Schematic of the four bispecific antibodies; two in an scFv format and two in a IgG-like CrossMAb format with knob-in-hole. The parental monoclonal antibodies that form the bispecific antibodies are colour-coded: blue, C135; orange, C144; green, C121; and purple, Fc region. **b**, All four constructs neutralize SARS-CoV-2 pseudovirus in vitro at sub-nanomolar concentrations: IC_{50} of 0.13, 0.04, 0.74 and 0.53 nM for CoV-X1, CoV-X2, CoV-scB1 and CoV-scB2, respectively. Normalized relative luminescence values (which correlate to infection) are reported versus antibody concentration, as previously detailed[15]. Mean of two replicates from one representative experiment is shown.
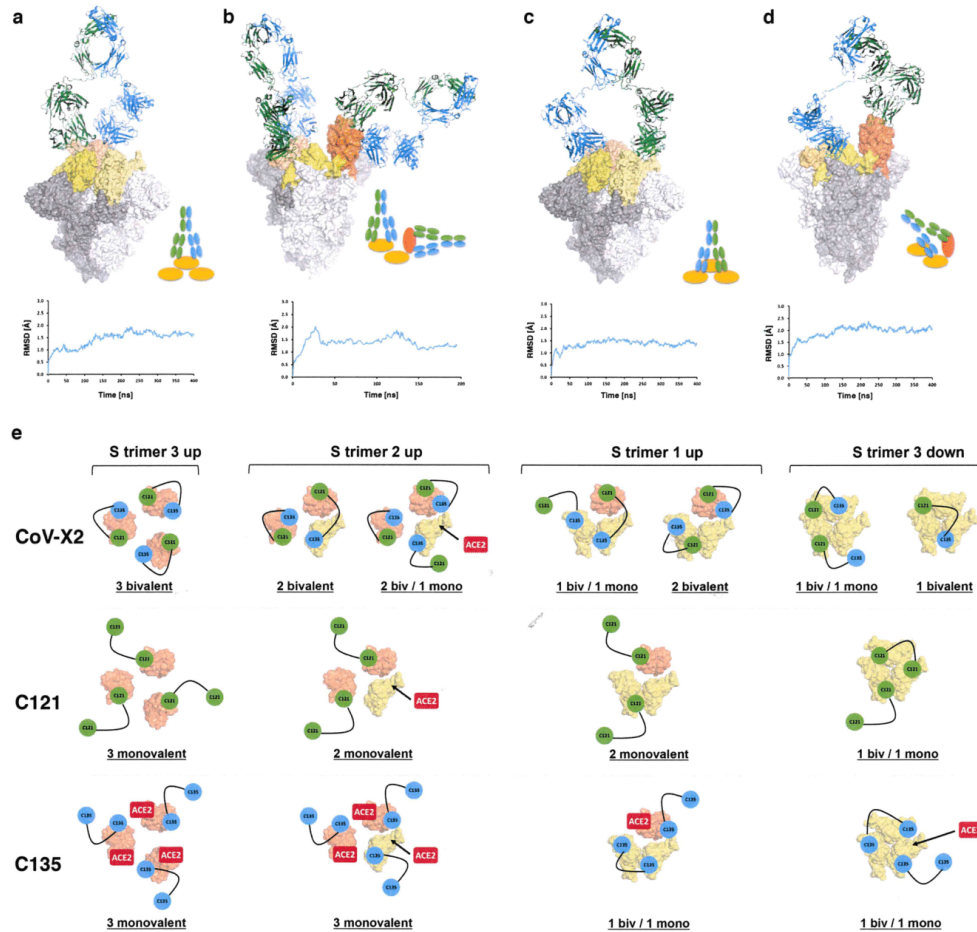
Extended Data Fig. 2 | CoV-X2 engages its epitopes on all RBD conformations on the S trimer. a–d, Molecular dynamics simulations of the complex between the CoV-X2 bispecific antibody and S trimers with RBD in either all down, all up or mixed up/down conformations show that CoV-X2 can engage a single RBD with both arms (**a**, **b**), two adjacent RBDs in the down conformation (**c**) and two RBDs in the up and down conformation (**b**, **d**). The complexes were subjected to up to 400 ns of fully atomistic molecular dynamics simulations to assess feasibility and stability of the bound conformations. Root-mean-squared deviations (r.m.s.d.) values are shown to indicate structural stability. S trimer is in shades of grey, RBDs in yellow (down conformation) and orange (up), the C121 and C135 moieties of CoV-X2 are in green and blue, respectively. **e**, Schematic of the computationally predicted binding modes of CoV-X2, C121 IgG and C135 IgG on the S trimer, coloured as in **a–d**. Antibodies are represented by connected circles; ACE2 is in red on the RBD if it can bind directly to a given conformation; and an arrow points to the RBD if ACE2 binding is achieved after an allowed switch to the up conformation. For example, in the three-up conformation (left), CoV-X2 can engage all the RBDs with bivalent binding, whereas C121 and C135 can achieve only monovalent binding. C135 binding does not prevent interaction with ACE2. The situation is similar in the other S conformations (two-up and one-down, two-down and one-up, and three-down); only the bispecific antibody achieves bivalent interaction and prevents ACE2 access in all conformations.

53

a

b

		CoV-X2	C121	C135
Spike protein	k_a [$\cdot 10^6$ M^{-1}s^{-1}]	0.69	1.33	0.62
	k_d [$\cdot 10^{-3}$s^{-1}]	0.12	0.17	0.13
	K_D [nM]	0.18	0.13	0.21
RBD	k_a [$\cdot 10^6$ M^{-1}s^{-1}]	0.58	1.30	0.20
	k_d [$\cdot 10^{-3}$s^{-1}]	1.35	3.15	0.12
	K_D [nM]	2.35	2.40	0.59
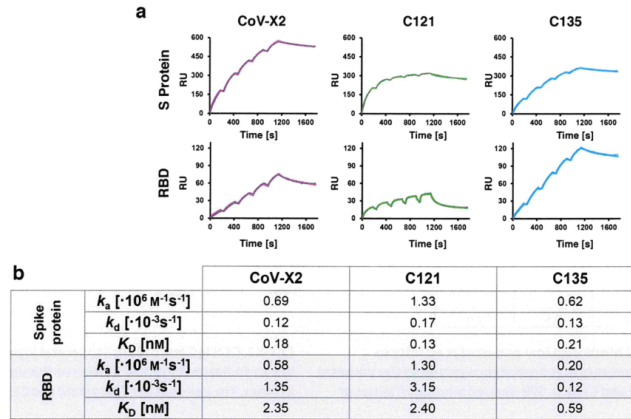
Extended Data Fig. 3 | CoV-X2 and its parental monoclonal antibodies bind recombinant, isolated RBD and S trimer with low nanomolar affinity.
a, Representative SPR traces from which the data in **b** were derived. **b**, Kinetic parameters for the binding of C121 IgG, C135 IgG and CoV-X2 to the S trimer and RBD.

a

	CoV-X2 K_D [пм]	C121 K_D [пм]	C135 K_D [пм]
Wild-type	2.3	2.4	0.6
N331A	2.1	0.3	0.4
E340K	1.4	1.9	<0.1
T345A	1.2	1.4	0.2
V367F	2.4	2.1	0.5
R408I	1.9	2.1	0.7
K417A	1.3	0.7	0.3
W436R	2.7	3.3	2.7
K444E	1.1	2.8	No binding
E484A/F486A	7.8	No binding	1.7
F490A	2.9	35	1.5
T500A	1.5	1.5	1.0
Y505A	1.4	1.6	0.6
R346T/A372T/I402V/ K417R/Q498H/H519K (Pangolin GD)	2.5	0.5	No binding
D614G	3.2	0.6	1.6

b
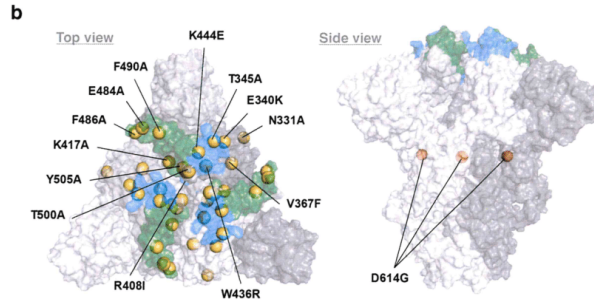
Extended Data Fig. 4 | CoV-X2 binds with low-nanomolar affinity to S mutants, including some mutants that are not recognized by the parental monoclonal antibodies C121 and C135. a, SPR-derived binding affinities of CoV-X2, C121 IgG and C135 IgG to several S trimer mutants. **b**, Mutations tested in **a** are indicated by yellow spheres on the surface representation of the S trimer. The epitopes of C121 (green) and C135 (blue) are shown.

a

Wild-type

K_D = 4.0 nM

B.1.1.7

K_D = 2.2 nM

B.1.351

K_D = 24 nM

b

N501
K417
E484

c

Wild-type
N501Y
K417N/E484K/
N501Y/R683G
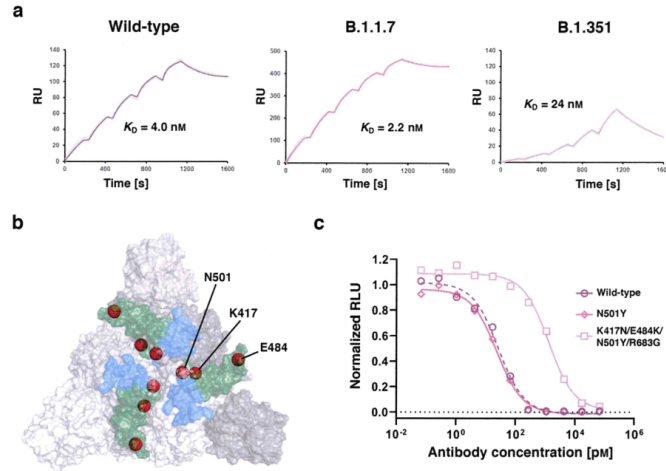
Extended Data Fig. 5 | Efficacy of CoV-X2 against the B.1.1.7 and B.1.351 variants. a, SPR traces showing binding of CoV-X2 to the RBD corresponding to wild-type SARS-CoV-2 and the B.1.1.7 and B.1.351 variants. **b**, Residues mutated in the variants are shown as red spheres on the surface representation of the S trimer. The epitopes of C121 (green) and C135 (blue) are shown.

c, Neutralization of SARS-CoV-2 pseudoviruses expressing wild-type S, S(N501Y) and S(K417N/E484K/N501Y/R683G) (corresponding to substitutions in the RBD in the B.1.351 variant (Fig. 1h)) by CoV-X2. Mean of two experiments is shown.

	Immobilized on the chip				
	RBD 150 nM	RBD 75 nM	RBD 15 nM	RBD 5 nM	
C135 Fab	1.0	1.2	1.7	1.7	k_a Normalized $k_a(X_i\,nM)/k_a(150\,nM)$
C135 IgG	1.0	1.1	1.0	0.9	
C121 IgG	1.0	1.1	3.1	2.6	
CoV-X2	1.0	1.4	1.8	2.0	
C135 Fab	1.0	1.1	1.0	0.9	k_d Normalized $k_d(X_i\,nM)/k_d(150\,nM)$
C135 IgG	1.0	2.3	9.1	18.8	
C121 IgG	1.0	3.2	9.0	18.5	
CoV-X2	1.0	1.3	3.2	5.7	
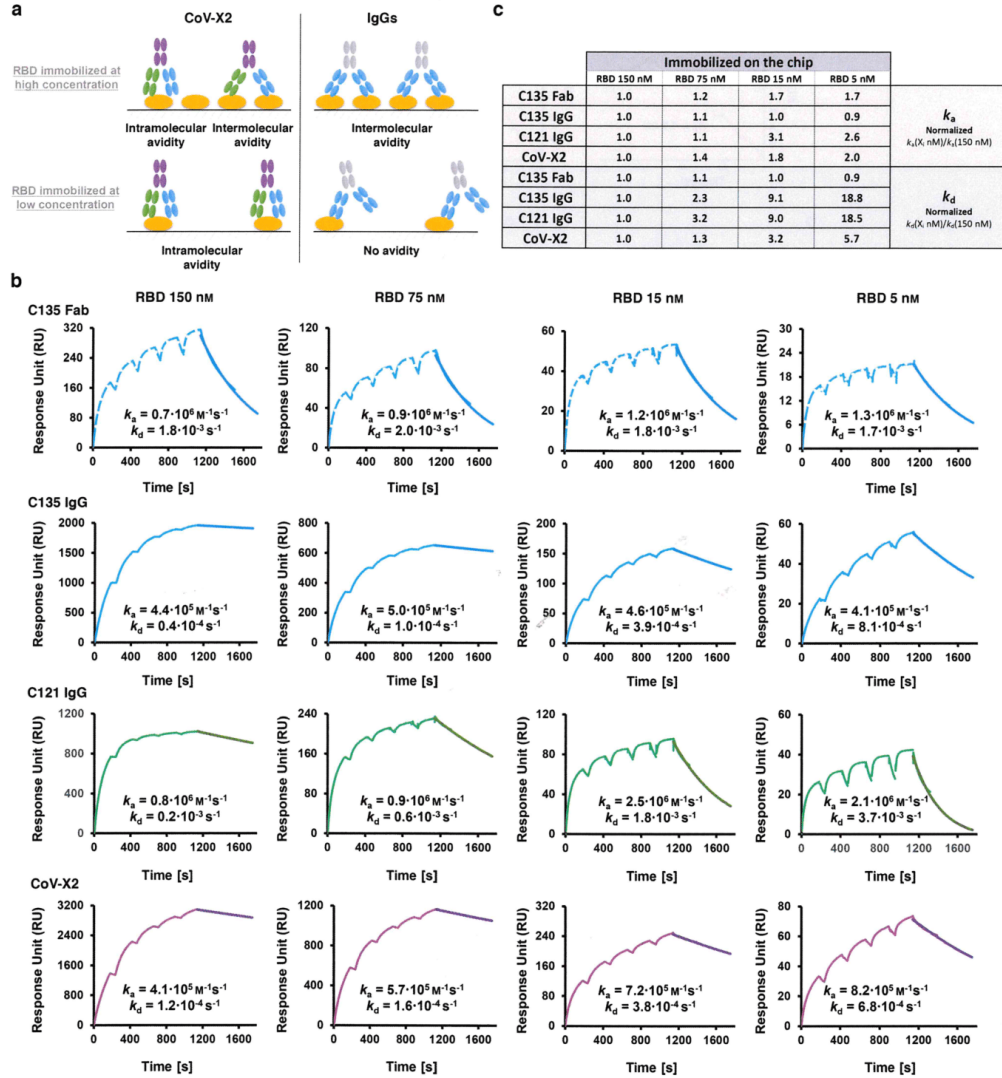
Extended Data Fig. 6 | SPR-based avidity assays confirm that CoV-X2 can engage bivalently on a single RBD. a, CoV-X2 and monoclonal IgGs (C121 or C135) have different binding modes that are available when high or low quantities of RBD are immobilized on the surface of the SPR chip. Monoclonal antibodies have avidity effects at high RBD concentrations owing to intermolecular binding (which results in a slower k_d), but not at low RBD concentrations as bivalent binding to a single RBD is impossible. By contrast, the bispecific antibody has avidity at both high and low concentrations, as bivalent binding to its two epitopes on a single RBD is possible. k_a is not affected by avidity. **b**, Experimental confirmation that CoV-X2 engages bivalently on a single RBD. SPR traces used to determine k_a and k_d of monoclonal antibodies, Fab and the bispecific antibody at different concentrations of immobilized RBD (Fig. 1d) are shown. **c**, Table summarizing the SPR results plotted in Fig. 1d. k_a and k_d were normalized against the values at the highest RBD concentration. k_a and Fab k_d were unaffected by the RBD concentration, as expected. k_d became faster for the monoclonal antibodies (loss of avidity) but less so for the bispecific antibody (avidity maintained owing to simultaneous binding to two sites on a single RBD).

57

Article

a

AAV-hACE2 transfer plasmid

AAV ssDNA genome (AAV-hACE2 vector)

b

| GC: | 5x10^10 | | | 1x10^11 | | |

hACE2 — 100 kDa

Histone H3 — 15 kDa

c

M 1 2 3 4 5 6 7 8 9 10 11 12 13 14

100 kDa
70 kDa
50 kDa

d

AAV-hACE2 dilutions	Ct1	Ct2	Ct3	Average Ct	No. of GC (log)	No. of GC	No. of GC counted from dilution
3125000x	25.63	25.68	25.70	25.67	3.526	3.36E+03	1.05E+10
625000x	24.70	24.60	24.70	24.67	3.874	7.48E+03	4.67E+09
125000x	22.84	22.69	22.83	22.79	4.526	3.36E+04	4.20E+09
25000x	20.65	20.61	20.86	20.71	5.248	1.77E+05	4.43E+09
5000x	18.33	18.43	18.25	18.34	6.070	1.18E+06	5.88E+09
1000x	16.15	16.00	16.05	16.07	6.858	7.21E+06	7.21E+09

	Average No. of GC/reaction	GC/µl	4x10^9 GC (µl)	PBS (to 40 µl)
AAV-hACE2	6.15E+09	2.46E+09	1.63	38.37

e

SARS-CoV-2 infection

D2   D5   D8   Control

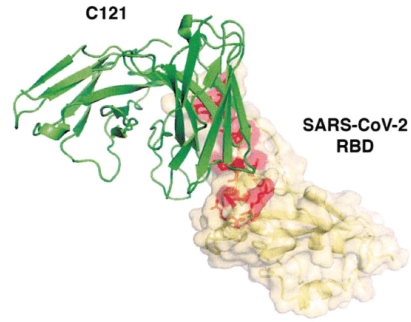**Extended Data Fig. 7** | See next page for caption.

58

Extended Data Fig. 7 | Generation of the AAV-hACE2-transduced mouse model of COVID-19. a, Diagram of the AAV-hACE2 plasmid and corresponding AAV vector. **b**, Western blot analysis detecting hACE2 expression in the lungs of one non-transduced control mouse (ctrl) and 12 mice transduced with two doses of AAV-hACE2 viral particles (5×10^{10} or 1×10^{11} genome copies (GC)). Lung tissue was collected 1, 2 or 4 weeks (w) after transduction. Histone H3 was used as control for quantification (bottom). Quantitative analysis represents normalized data from membrane images (top), and was performed using ImageJ. Representative data from two independent experiments are shown. **c**, Preparation of concentrated AAV-hACE2. AAV-hACE2 plasmid was co-transfected with pHelper and AAV Rep/Cap 2/9n vectors into 293AAV cells (Methods). To increase viral titres, viral particles from both cell lysate and PEG-precipitated growth medium were ultracentrifuged in a discontinuous iodixanol gradient. The silver-stained SDS–PAGE gel shows 14 consecutive fractions: fractions 1–9 represent enriched AAV fractions used for experiments, and fractions 10–14 are contaminated with proteinaceous cell debris. Iodixanol was chosen as a density gradient medium owing to its low toxicity in vivo and its easy removal by ultrafiltration. M, protein marker; *AAV capsid proteins VP1, VP2, and VP3. Representative data from two independent experiments are shown. **d**, The amount of AAV particles was estimated by RT–qPCR. The number of genome copies expressed as log was calculated from a standard curve. From one 15-cm² dish, 75 µl with 2.0×10^{12} genome copies per ml were prepared, which is sufficient for hACE2 humanization of 37 mice. **e**, Kinetic of lung histopathology in SARS-CoV-2-infected ACE2-humanized mice. H&E-stained sections showed inflammatory infiltrates composed of lymphocytes, macrophages, neutrophils and fibroblasts replacing the alveoli. The size of the affected areas increased over time (area of diffuse alveolar damage: control <5–10%, 2 dpi <10–30%, 5 dpi 20–80% and 8 dpi 50–90%). Alveolar septa were thickened in areas that were close to infiltrates. In samples collected at 5 and 8 dpi, an increased number of activated macrophages with foamy cytoplasm (black arrowheads) was seen. AAV-hACE2-transduced, SARS-CoV-2-uninfected mice were used as control and showed no noticeable pathology. Each image is representative of two separate experiments ($n = 3$ to 5 mice per group).

Article

a

wt	mutant	Strain Name	wt	mutant	Strain Name
Arg403	Lys	/USA/VA-DCLS-0630/2020 /USA/VA-DCLS-0439/2020 /AUS/VIC1787/2020			
Lys444	Asn	/AUS/VIC4515/2020	Gly485	Arg	/AUS/VIC1829/2020 /AUS/VIC1960/2020 /AUS/VIC1693/2020 /AUS/VIC1660/2020 /AUS/VIC1683/2020 /AUS/VIC1588/2020 /AUS/VIC1565/2020 /AUS/VIC1611/2020 /AUS/VIC1812/2020 /AUS/VIC2023/2020
Gly446	Asp	/USA/FL-BPHL-2211/2020			
	Val	/USA/MN-MDH-1430/2020 /AUS/VIC913/2020 /AUS/VIC6087/2020 /AUS/VIC9542/2020			
	Gln	/USA/VA-DCLS-1404/2020	Cys488	Arg	/RN/COVID19-IRVSH4/2020
Leu452	Arg	/USA/CA-CZB-12872/2020	Phe490	Leu	/AUS/VIC10024/2020 /AUS/VIC766/2020
	Met	/BHR/340798279_S5_L001/2020 /USA/CA-CZB-1043/2020 /IND/906/2020		Val	/USA/MA-MGH-01415/2020
Leu455	Phe	/AUS/VIC10121/2020 /AUS/VIC5196/2020	Pro491	His	/IND/906/2020
			Gln493	Leu	/USA/WI-UW-371/2020
Val483	Ala	/USA/WA-UW-6527/2020 /USA/WA-UW-1587/2020 /USA/WA-RML-2/2020 /USA/WA-RML-6/2020 /USA/WA-RML-5/2020 /USA/UT-03764/2020	Ser494	Pro	/USA/CA-CZB-4047/2020 /USA/CA-CZB-11677/2020 /USA/CA-CZB-6994/2020 /USA/CA-CZB-11010/2020 /USA/MI-MDHHS-SC20047/2020 /USA/CA-CZB-12810/2020 /AUS/VIC9505/2020
	Phe	/AUS/VIC2139/2020 /USA/MA-UW-629/2020		Leu	/AUS/VIC5057/2020
	Gln	/USA/UT-UPHL-2009538/2020 /IND/GBRC278a/2020 /USA/SEARCH-1462-SAN/2020			
Glu484	Lys	/USA/UT-QDX-1869/2020 /BHR/340859913_511/2020 /USA/IL-UW-379/2020			
	Ala	/USA/VA-DCLS-1615/2020			

b

wt	mutant	Strain Name
Phe342	Leu	England/01, 1/29
Ala344	Ser	/USA/WA-S2278/2020 /USA/WA-S2530/2020
	Thr	/IND/GBRC431a/2020
	Val	/AUS/VIC10958/2020
Thr345	Ser	/USA/WA-S1049/2020
	Ile	/PER/covper051/2020
Arg346	Thr	/IND/GBRC333/2020
Trp436	Thr	/IND/GBRC333/2020
Asn439	Lys	/USA/IL-UW/799/2020
Asn440	Lys	/HKG/Case5138/2020
Leu441	Ile	/USA/FL-BPHL-0297/2020
Asn450	Lys	/IND/906/2020
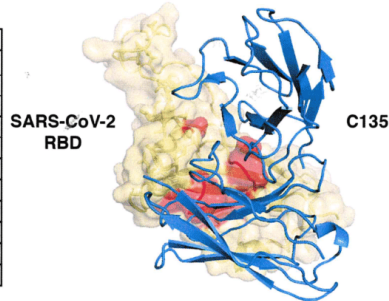
Extended Data Fig. 8 | Natural SARS-CoV-2 variants in the C121 and C135 epitopes. a, **b**, Summary of naturally occurring mutations in the C121 (**a**) or C135 (**b**) epitopes that have been reported in circulating SARS-CoV-2 (as of 1 January 2021). The location of the mutated residues is shown in red on the RBD structure. The C121 and C135 variable regions are in green and blue, respectively (PDB 7K8X and 7K8Z, respectively). All the variants were taken from the ViPR database (https://www.viprbrc.org/).

Extended Data Table 1 | Summary of the *P* values for the mouse protection experiment

	C121	C135	CoV-X2	Isotype control
C121	–	P<0.0001	P<0.0001	P<0.01
C135	P<0.0001	—	P>0.05	P<0.0001
CoV-X2	P<0.0001	P>0.05	—	P<0.0001
Isotype control	P<0.01	P<0.0001	P<0.0001	—

Statistical comparison of body weight differences in mice treated with the individual monoclonal antibodies (C121 or C135), the CoV-X2 bispecific antibody or isotype control at 8 dpi (data shown in Fig. 2e). *P* values were determined with a one-way analysis of variance (ANOVA). Comparison of the entire curves (Fig. 2e) by the one-sample Wilcoxon test or by ANOVA followed by Turkey–Kramer post test reveals that the isotype-control-treated group is statistically different from any one of the other groups (CoV-X2, *P* = 0.0159; C135, *P* = 0.0043; and C121, *P* = 0.0010).

4.2   Part 1b: COVID-19 mouse models and their AAV-based transient alternative

This comprehensive review article provides a detailed comparison of humanized transgenic models and transient (sensitized) humanized mouse models. The primary goal is to highlight the advantages and limitations of these models while emphasizing their unique characteristics. Additionally, the review discusses factors influencing COVID-19 severity, including obesity, diabetes, age, and hypertension. Lastly, it situates our AAV-hACE2 transient mouse model within the broader context of other available transgenic and transiently humanized models.

Check for updates

Advances in Modelling COVID-19 in Animals

Petr Nickl[1,2†], Miles Joseph Raishbrook[1,2†], Linn Amanda Syding[1†] and Radislav Sedlacek[1,2]*

[1]*Laboratory of Transgenic Models of Diseases, Institute of Molecular Genetics of the CAS, v.v.i, Vestec, Czechia,* [2]*Czech Centre for Phenogenomics, Institute of Molecular Genetics of the CAS, v.v.i, Vestec, Czechia*

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV2) is a positive-sense-single stranded RNA virus and the cause of the coronavirus disease 2019 (COVID-19). The World Health Organisation has confirmed over 250 million cases with over 5.1 million deaths as a result of this pandemic since December 2019. A global outbreak of such intensity and perseverance is due to the novelty of SARS-CoV2 virus, meaning humans lack any pre-existing immunity to the virus. Humanised animal models, from rodents to primates, simulating SARS-CoV2 transmission, cell entry and immune defence in humans have already been crucial to boost understanding of its molecular mechanisms of infection, reveal at-risk populations, and study the pathophysiology *in vivo.* Focus is now turning towards using this knowledge to create effective vaccines and therapeutic agents, as well as optimise their safety for translatable use in humans. SARS-CoV2 possesses remarkable adaptability and rapid mutagenic capabilities thus exploiting innovative animal models will be pivotal to outmanoeuvre it during this pandemic. In this review, we summarise all generated SARS-CoV2-related animal models to date, evaluate their suitability for COVID-19 research, and address the current and future state of the importance of animal models in this field.

Keywords: COVID-19, mouse, model, SARS-CoV2, sensitised, humanized, mice

1 COVID-19 ORIGIN AND SARS-COV2 TRANSMISSION

A pandemic is defined as a disease that is prevalent in an entire country or the world, and thus is undoubtedly the correct term for the coronavirus disease 2019 (COVID-19) outbreak. COVID-19 is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV2) which has been pinpointed to have originated from Wuhan, China in December 2019 and has since then spread over all continents including Antarctica (Triggle et al., 2021). Before the COVID-19 outburst there were already two identified and relatively well-known human coronaviruses causing severe respiratory pneumonia namely, SARS-CoV and MERS-CoV. They both originate from bats but spilled over to intermediate hosts namely, civets and dromedary camels, respectively. The origin of SARS-CoV2 is also suggested, based on its sequence similarity to the SARS-CoV, to have originated from bats and later spilled over to an animal reservoir, however it is not yet confirmed (Forni et al., 2017). Bats are and continue to be a copious source for novel viral sequences (Jiang et al., 2022). The bat species are among one of the oldest mammals and they exhibit great diversity and are widely spread across the globe (X. M. Zhang et al., 1992). Cross-species mixing between different kinds of bats has facilitated a maintenance of less discriminatory viruses capable of infecting a broader variety of hosts. Bats are thus a carrier of a pool of viruses able to perform inter-species transmission, which has been a reason for concern long before the COVID-19 outbreak (Calisher et al., 2006). SARS-CoV2 is a pneumotropic virus that mainly spreads through respiratory secretions like coughing and sneezing. The transmission may also occur *via* contaminated surfaces where the virus can

63

FIGURE 1 | ACE2-mediated entry of SARS-COV2 into a cell **(A)** The SARS-CoV2 virion consists of structural proteins, namely spike (S), envelope (E), membrane (M), nucleocapsid (N) The positive-sense, single-stranded RNA genome (+ssRNA) is bound by N in a beads-on-a-string formation. **(B)** SARS-CoV2 binds to the cell *via* S interaction with the host's ACE2 receptor. Entry to the host cell either goes through receptor-mediated endocytosis where fusion is potentiated by the cleavage of S2 by Cathepsin L or by cell surface fusion *via* the TMPRSS2 serine protease. Following fusion, the virion is uncoated and the viral genome released. Created with Biorender.com.

survive up to 6 days, making preventive measures such as surface disinfection, hand hygiene and masks important in combating transmission (Leclerc et al., 2020).

Once infected, COVID-19 manifests in a large variety of symptoms however, the most common ones include fever, sore throat, fatigue, cough, dyspnoea and immune system dysregulation, often ending with cytokine storm (Ragab et al., 2020). SARS-CoV2 is not only capable of affecting the respiratory system by pulmonary infiltration and inflammation but can spread to multiple organ systems. For the majority of people the disease symptoms are mild and the infection runs its course without any medical intervention but for approximately 5–10%, it severely affects the fitness of the individual and for another 2% it has a mortal outcome (Gavriatopoulou et al., 2021).

The SARS-CoV2 consists of approximately 29.9kB of single-stranded, non-segmented, positive-sense RNA.(Triggle et al., 2021). The genome is composed by 13–15 open reading frames largely resembling the make-up of MERS-CoV and SARS-CoV. The genome contains 11 protein coding genes with ultimate expression of 12 expressed proteins (Lu et al., 2020).

Structurally, it consists of four proteins namely, spike (S), envelope (E), membrane (M) and nucleocapsid (N) (**Figure 1A**). These proteins play important parts in entry, fusion and replication in the host cells. The non-structural have roles imperative to viral pathogenesis by regulating early transcription, helicase activity, gene transactivation and countering antiviral response (J Alsaadi and Jones, 2019; Tang et al., 2020).

64

The spike glycoprotein plays a pivotal role in the pathogenesis of SARS-CoV2 as it pivotal for the entry into the host cell. It is assembled as a homotrimer and inserted in multiple copies into the virus membrane, giving the virus a crown-like appearance, thus its name coronavirus (Jackson et al., 2022). It consists of two functional subunits, S1 and S2, that both part take in the entry of the virus. The S1 subunit has a receptor-binding domain (RBD) and is responsible for anchoring the host cell upon binding between the RBD and the human angiotensin-converting enzyme 2 (hACE2), thus stabilizing the virus (Hoffmann et al., 2020). Once the RBD region of the S1 subunit binds to the hACE2, the virus enters the host's endosomes *via* ligand-mediated endocytosis or membrane-fusion. Once bound to the ACE2, the S protein undergoes conformational changes, which are important to therapeutically limit its infection cycle (Wrapp et al., 2020). Although several mutations have been found in the RBD of the S1 subunit, its affinity to and interaction with the hACE2 is preserved in most species however not in mouse (Chan et al., 2020; Wrapp et al., 2020). The S2 subunit functions as a fusion protein between the virus and the host cell membrane. The S2 exhibit three different conformational changes during the process namely, i) native state before fusion, ii) intermediate state and iii) post fusion hairpin state (Qing and Gallagher, 2020; Walls et al., 2020). Finally the S protein is cleaved either by the host cell surface serine protease TMPRSS or by host's Cathepsin L in the endosomal compartment at the S2′ cleavage site (**Figure 1B**; Simmons et al., 2005). The cleavage releases a fusion peptide, which initiates the fusion pore formation. Once the pore expands and the cell membranes of both the virion and the host are combined, the viral genome can be released in to the cytoplasm. The cell membrane or endosomal fusion, represent the two different modes of entry for the viral genome to be released.

The N protein, composed by two separate domains, is present in the nucleocapsid complex that tightly binds the RNA genome of the virus. Both the N-terminal and C-terminal domain can bind to RNA but is more efficient when both bind simultaneously (Chang et al., 2006). The N protein bind the viral RNA genome in a beads-on-a-string conformation. The ribonucleotide protein (RNP) complex is subsequently packaged in to viral particles enveloped by a fatty lipid bi-layer (Fehr and Perlman, 2015).

The envelope protein is a relatively small protein that plays a substantial role in viral assembly. The protein assemble in to the host membrane forming protein-lipid pores referred to as viroporins. The envelope protein is highly conserved between SARS-CoV and SARS-CoV2 (Fehr and Perlman, 2015).

SARS-CoV2's membrane protein is the most abundant structural protein and is a transmembrane with a short NH2 terminal on the outside and a long cytoplasmic COOH terminus. Completion of viral assembly is potentiated partly by the binding between M proteins and N proteins leading to a stabilization of the N-Protein and RNA complex internally (Thomas, 2020).

2 INFECTION ROUTE AND ACE2 FUNCTION

The primary route of entry for the SARS-CoV2 is the upper respiratory tract. The virus gains access to the host cells by binding to the ACE2 receptors and subsequently introduced in to the cytoplasm *via* receptor-mediated endocytosis. The virus particles then goes through uncoating. The RNA and proteins needed for translation are released followed by transcription and assembly, finally the viral loads are shed thus completing the viral replication cycle (Jiang et al., 2020). As the virus sheds, the newly replicated and released particles bind upon another host cell and the cycle starts again. The ACE2 receptor is a carboxypeptidase consisting of 805 amino acids that removes a single amino acid from the C terminus of its substrates (Turner and Hooper, 2002). The ACE2 receptors are expressed in alveolar epithelial cells and capillary endothelial cells that are abundant in organs such as the lungs, kidneys, brain and gut hence explaining the multisystem infection found in a substantial amount of patients (Samavati and Uhal, 2020). The physiological role of ACE2 in humans is to convert angiotensin I and II to angiotensin 1–9 and angiotensin 1–7, respectively. This is one of the steps making up the Renin Angiotensin Aldosterone System (RAAS) a system, which functions to elevate blood volume and arterial tone *via* sodium and water reabsorption and vascular tone (Nehme et al., 2019). Infection results in a decrease of physiologically available ACE2 receptors thus disrupts the RAAS system, leading to potential downstream complications such as inflammation and circulatory dysfunction (Guo et al., 2020).

3 TRANSLATIONAL STUDIES: IMPORTANCE OF MOUSE MODELS

Since the start of the COVID-19 pandemic, we have gained substantial knowledge about the SARS-CoV2 virus in terms of its genetic make-up, transmission, infection and pathogenesis. This allows us to develop therapeutic agents to combat it. However, to perform scientifically sound and reliable research it is of the utmost importance to work with an appropriate model organism for *in vivo* study. The laboratory mouse is the most used animal in medical research as they are inexpensive, easy to handle, are genetically very similar to humans and can be genetically modified relatively easy (Sellers, 2017). They are often present as inbred strains, making it a highly controlled system, which is desirable in medical research. The mouse as an organism for translational research in COVID-19 medical research is however not well suited for COVID-19 as the ACE2 receptor of the mouse is not efficiently bound by the SARS-CoV2 virus, thus rendering the mouse immune to severe infection. This seemingly huge barrier has been surpassed by the generation of various modified mouse models capable of infection (Jia et al., 2020), as exemplified in the text below.

The COVID-19 outbreak pointed out a desperate need for relevant animal models for SARS-CoV2 research. As mentioned above wild type mouse cells and tissues are not very susceptible to SARS-CoV2 due to lack of human ACE2 specifically. Basically, mouse Ace2 does not bind the virus efficiently enough to mediate cell entry. To overcome this obstacle and study COVID-19 in mouse models, researchers have developed several approaches such as "*murinisation*" of SARS-CoV2 (Dinnon et al., 2020; Gu

65

TABLE 1 | Overview of COVID-19 mouse models and their characteristics.

Transgenic mouse model/background	Promoter/tissue	SARS-CoV2 dose/Most affected tissues/Symptoms/Lethality (intranasal route-IN, intravascular-IV)	References
Krt18-hACE2 (C57BL/6	Epithelial cell cytokeratin-18 promoter, epithelial cells	2.5×10^4 PFU (IN)/lung, kidney, brain, heart, spleen/severe interstitial pneumonia/lethal	McCray et al. (2007), Winkler et al., (2020)
HFH4-hACE2 (C3H, C57BL/6)	HFH4/FOXJ1 - lung ciliated epithelial cell-specific promoter, predominantly expressed in lung (also detected in brain, liver, kidney and gut)	3×10^4 TCID$_{50}$ (IN)/lung, heart, eye, brain/ severe interstitial pneumonia/lethal	Menachery et al. (2016), Jiang et al. (2020)
pCAGGS-hACE2 (C57BL/6 or BALB/c)	Cytomegalovirus enhancer with chicken β-actin promoter/universal expression	2×10^5 or 10^3 TCID$_{50}$ (IN) of SARS-CoV/ lungs, brain/acute wasting syndrome/lethal	Tseng et al. (2007)
Ace2-hACE2 Ace2-hACE2-IRES-tdTomato (ICR,C57Bl/6	Murine angiotensin converting enzyme 2/ intestine, brain, heart, kidney	10^5 TCID$_{50}$ (IN)/lung, intestine, brain/ moderate interstitial pneumonia/non-lethal	Bao et al. (2020a), Sun S.-H. et al. (2020), Yang et al. (2007b)
hACE2(LoxP-STOP) (C57Bl/6J)	Cytomegalovirus enhancer with chicken β-actin promoter/conditioned expression of hACE2-IRES-eGFP cassette	4.5 l g FFU (IN)/lung, brain/dramatic weight loss and rapid mortality/lethal (ubiquitous expression)	Bruter et al. (2021), Dolskiy et al. (2022)
Rosa26-chACE2 (C57Bl/6N)	Cytomegalovirus enhancer with chicken β-actin promoter/conditioned expression	2×10^3 PFU (IN)/not characterized/weight loss and rapid mortality/lethal (ubiquitous expression)	Czech Centre for Phenogenomics (2021)
Sensitised mouse models	**Promoter/tissue**	**SARS-CoV2 dose/Most affected tissues/Symptoms/Lethality (intranasal route-IN, intravascular-IV)**	**References**
AdV-hACE2/AdV-hACE2-GFP (BALB/c; C57BL/6J; Rag1$^{-/-}$ C57BL/6, Stat1$^{-/-}$ C57BL/6; DBA/2J; AG129)	Cytomegalovirus promoter, lung	10^5 FFU (IN); 10^5 PFU(IN, IV)/lung, heart, brain, liver, spleen/weight loss/non-lethal	Hassan et al. (2020), Sun J. et al. (2020)
AAV-hACE2 (C57BL/6J,B6(Cg) Ifnar1tm1.2Ees/J(Ifnar1$^{-/-}$); C57BL/6NCrl	Cytomegalovirus promoter/lung*	3×10^7 PFU/ml (IN); 1×10^4 PFU (IN)/lung (other organs not characterized)/weight loss/ non-lethal (*note: Localization of AAV-hACE2 expression is dependent on route of AAV application and used AAV serotype.)	Israelow et al. (2020), De Gasparo et al. (2021)
Lenti-hACE2 (C57BL/6J, IFNAR–/–)	Elongation factor 1 alpha promoter/lung	2×10^5 pfu (IN); 10^5 CCID$_{50}$ per mouse (IN)/ lung (inflammatory response)/mild symptoms of the COVID-19 disease, weight loss/non-lethal	Rawle et al. (2021), Katzman et al. (2022)

et al., 2020) or humanisation of mouse models (McCray et al., 2007; Tseng et al., 2007; Menachery et al., 2016). Alternatively they used different animal models which are sensitive to known SARS-CoV2 variants, such as hamsters, ferrets or non-human primates (Enkirch and von Messling, 2015; Finch et al., 2020; Munster et al., 2020; Rockx et al., 2020; Gruber et al., 2021).

Several transgenic mouse models have been developed and used in COVID-19 research to overcome limits of mouse Ace2 and generate inexpensive models with high-throughput study potential. The models are based on ubiquitous or cell/tissue specific expression of human ACE2, a protein well-known for its importance in SARS-CoV2 entry in the cell.
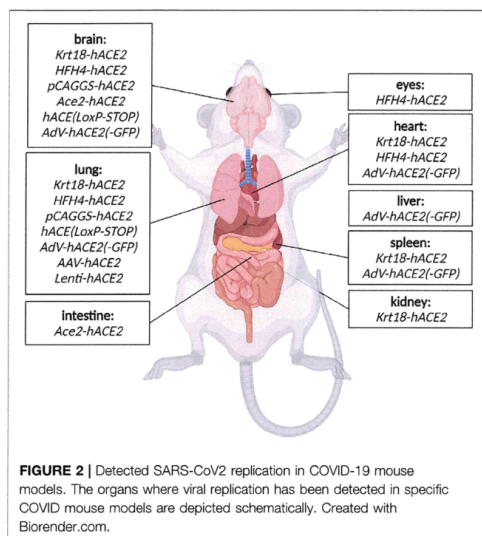
3.1 K18-hACE2 Model

This mouse model expresses hACE2 under control of the epithelial cell cytokeratin-18 promoter, expressed in subset of epithelial cells (**Table 1**). Even though the model has been generated to study SARS-CoV during 2003 epidemics, it remains relevant for SARS-CoV2 research too (McCray et al.,

2007). K18-hACE2 model is susceptible to both strains of SARS-CoV, one and 2. In the context of COVID-19 research, the model responds to the infection by progressive weight loss, high viral titres in the lung at the beginning of infection, and with progressing infection increasing viral titres in the brain and gut. Other less severely affected organs are heart, kidney and spleen (**Figure 2**). (McCray et al., 2007; Rathnasinghe et al., 2020).

3.2 HFH4-hACE2 Model

The model expresses hACE2 under lung ciliated epithelial cell-specific promoter (HFH4/FOXJ1), which was supposed to drive lung-specific hACE2 expression. However, detailed characterization of the model revealed a moderate hACE2 expression in other tissues such as the brain, eye, heart, liver, kidney and gut (**Figure 2**). The model is highly responsive to SARS-CoV2 infection with main replication of the virus in lungs, eyes, heart and brain accompanied with severe symptoms such as interstitial pneumonia sometimes succumbed to lethal encephalitis (**Table 1**) (Menachery et al., 2016; Jiang et al., 2020).

66

FIGURE 2 | Detected SARS-CoV2 replication in COVID-19 mouse models. The organs where viral replication has been detected in specific COVID mouse models are depicted schematically. Created with Biorender.com.

3.3 pCAGG-hACE2 Model

A model generated with multiple random integrations of pCAG-hACE2 cassette throughout the genome. Cytomegalovirus enhancer with chicken β-actin promoter (CAG) allows ubiquitous and constant expression of hACE2 in all tissues but mainly in lung, brain, heart and kidney (**Table 1**). This model is highly susceptible to SARS-CoV1 and 2 after intranasal application. The infection starts with initial exponential growth of viral titre in lungs and continues with gradual transmission to the brain. Slight presence of the virus was also observed in heart, kidney, spleen and small intestine (**Figure 2**). Of note, the lethal titre of the virus for the pCAGG-hACE2 model (2×10^2 to 2×10^4 TCID$_{50}$) is lower than in case of K18-hACE2 model (10^4 to 10^5 TCID$_{50}$) This fact might be connected to multiple insertion of pCAGG-hACE2 cassette in the genome in combination with ubiquitous and strong expression of hACE2 (Tseng et al., 2007; Asaka et al., 2021).

3.4 Ace2-hACE2 Model

Two major Ace2-hACE2 models have been generated to more closely mimic expression pattern of Ace2. The first model by was generated by random integration of a hACE2 cDNA under control of Ace2 promoter (Yang X. H. et al., 2007). The model indeed recapitulates endogenous expression of Ace2 in tissues such as lung, kidney, heart and intestines; the model is responsive to SARS-CoV2 infection with the major impact on lung tissue (Bao et al., 2020b).

The second Ace2-hACE2 model, generated by Sun S.-H. et al. (2020), is based on replacement of Ace2 coding sequence with hACE2 and tdTomato cDNA (**Table 1**). Therefore, the expression of transgenic cassette hACE2-IRES-tdTomato is

under control of endogenous Ace2 promoter and present only in one or two copies depending on zygosity. Despite the lack of clinical symptoms or elevated mortality, this model responds to SARS-CoV2 infection by interstitial pneumonia of distinct scale depending on age. Sun's group also points out different abundance of hACE2 throughout-tissues in human (kidney, heart, oesophagus, bladder, ileum) and hACE2 in their model (liver, spleen, small intestine, ovary, and brain), however without further explanation. Furthermore, the group identified brain, lung and trachea as the main tissues of SARS-CoV2 replication (**Figure 2**; Sun S.-H. et al., 2020).

3.5 hACE2(LoxP-STOP) Model

hACE2 (LoxP-STOP) also termed TgCAGLoxPStopACE2GFP is a model generated by random integration of a loxP-CRE dependent cassette under the CAG promoter (**Table 1**). In the presence of Cre recombinase, the STOP cassette is removed and expression of hACE2 cDNA and eGFP is turned on. This model allows for conditioned, tissue-specific and traceable expression of hACE2-IRES-GFP transgene (Bruter et al., 2021). Dolskiy and collective have tested two inducible and ubiquitously expressed Cre-ERT2 drivers (UBC-ACE2 and Rosa26-ACE2) to promote conditioned hACE2 expression. In this case, the most severely affected organs were lung and brain (**Figure 2**). Their results further suggested that severity and infection progress is dependent on the particular Cre driver, more specifically on its expression potency. Furthermore, relatively recent changes in renin-angiotensin system due to hACE2 overexpression can be another factor influencing response to the infection (Dolskiy et al., 2022).

3.6 Rosa26-chACE2

A model similar to the previous one, but a CAG-LoxP-STOP-LoxP-hACE2 cassette is inserted in Rosa26 locus in a site-specific manner (**Table 1**). Therefore, transgene copy number depends on zygosity. The model has not been validated yet through SARS-CoV2 infection. It is available at Czech Centre for Phenogenomics and will be soon available *via* European Mouse Mutant Archive (EMMA) (Czech Centre for Phenogenomics, 2021).

Of note, transgenic models have an important role in SARS-CoV2 research. However, their ectopic expression of ACE2 protein, specifically in case of K18-hACE2, HFH4-hACE2, and pCAGG-hACE2 models may lead to different response, development and impact of the infection. This fact to some extend limits translatability of gathered data to clinical practise (Shou et al., 2021). Therefore, ACE2 under control of endogenous Ace2 promoter or conditional expression might provide more precise understanding of systemic or tissue-specific importance of ACE2 in the context of COVID-19.

3.7 Sensitised Mouse Models

In order to circumvent the desperate need for COVID-19 mouse models in the peak of pandemics, researchers focused on development of alternative SARS-CoV2- sensitive mouse models. Paradoxically, a rapid generation of such models was mediated by viruses. Inhalation or intranasal application of a viral

67

vector carrying hACE2 gene under strong promoter may lead to humanisation of upper and lower respiratory tracts. This approach allows fast, affordable and versatile generation of a sensitive model in various mouse strains and genetic backgrounds. It has been shown that the most suitable viral vectors for rapid humanisation happen to be adeno-associated virus, adenovirus, and lentivirus.

3.8 AAV-hACE2

Two independent groups have used Adeno-Associated vector of serotype 9 to deliver a cassette with hACE2 under control of CMV to the lung (**Table 1**; Israelow et al. (2020) have used commercially available AAV-CMV-hACE2 plasmid for AAV production and applied the vector virus *via* injection into the trachea. The De Gasparo's group assembled the AAV-CMV-hACE2 plasmid by subcloning hACE2 cDNA isolated form HEK293 cells and the vector was administered with forced inhalation into the lung and upper respiratory system. Both groups confirmed functionality of AAV-mediated humanisation where treated mice became susceptible to SARS-CoV2 infection accompanied with progressive inflammatory immune response in lung (**Figure 1**; Israelow et al., 2020; De Gasparo et al., 2021). In addition to establishing a new sensitized model, Israelow and collective focused on deciphering the role of type I interferon during SARS-CoV2 infection. Whereas, De Gasparo and collective tested bispecific antibodies that reduced SARS-CoV2 infection and weight loss associated with ongoing virus infection. Humanisation with AAV offers rapid, adaptable mouse model with long-term transgene expression and low immunogenicity which is crucial for immunological studies (De Gasparo et al., 2021; Kovacech et al., 2022).

3.9 AdV-hACE2

Replication defective Adenovirus encoding hACE2 (AdV-hACE2) was used to humanise several mouse strains in order to overcome unavailability of transgenic models (**Table 1**). The vector is delivered intranasally and it is capable to sensitise lung tissue for SARS-CoV2 entry and replication. In other organs, low levels of SARS-CoV2 replication was also identified, such as heart, spleen, brain and liver (**Figure 2**). Sensitised models suffer from weight loss, develop lung pathologies and respond positively to treatment with neutralising antibodies. However, the model has limitations in the form of bronchial inflammation associated with AdV delivery (Hassan et al., 2020).

3.10 Lenti-hACE2

Lentiviral vectors can be also used for sensitising a mouse to SARS-CoV2. Two independent publications describe utility of a lentiviral vector encoding hACE2 and its ability to avoid significant immune response in lung tissue before SARS-CoV2 exposure (**Table 1**). The advantage of lentiviral systems is their integrative character, with possibly stable long-term expression allowing re-infection studies in the sensitised mice. Both publications emphasize the role of IFNAR1 depletion and its impact on SARS-CoV2 progression in sensitised models. However, the collectives also point out the presence of mild COVID-19 symptoms in the models, probably due to relatively

low expression of hACE2 by lentivirus (Rawle et al., 2021; Katzman et al., 2022).

Transgenic mouse models, expressing hACE2, represent convenient systems for large-scale, rapid (compared to other animal models), and relatively inexpensive SARS-CoV2 research. However, their availability during pandemics has been limited and their expansion in larger cohorts is time-consuming and expensive. Furthermore, distinct transgenic models differ in their response to infection, some suffer from lethal neuroinvasion, some show only mild symptoms. In general, variability of these models is significant, and no universal transgenic model has been established yet (Yang XH. et al., 2007; Yang XH. et al., 2007; McCray et al., 2007; Tseng et al., 2007; Menachery et al., 2016; Jiang et al., 2020; Sun J. et al. (2020) Bruter et al., 2021; Dolskiy et al., 2022).

In contrast with transgenic models stand virus-sensitised models, which can be generated on wide variety of genetic backgrounds and genotypes in relatively large scale and short-time. Sensitised models often do not develop severe disease mainly due to absence of neuroinvasion, but their symptoms and impact on lung tissue resembles pathology in COVID-19 patients. Moreover, the distribution and scale of hACE2 expression varies with tropism of a used viral vector or promoter. Importantly, use of viral vectors may be associated with a risk of potential inflammation leading to interference with subsequent SARS-CoV2 infection (Hassan et al., 2020; Israelow et al., 2020; De Gasparo et al., 2021; Rawle et al., 2021; Katzman et al., 2022).

3.11 Other Animal Models

Alternatives to mouse models are other animals that are naturally susceptible to SARS-CoV2, such as hamsters (*Mesocricetus auratus*, *Phodopus roborovskii*, *Cricetulus griseus*), ferrets (*Mustela putorius furo*) minks (*Neovison vison*) (Shuai et al., 2021)and non-human primates (*Macaca mulatta*, *Macaca fascicularis*, *Chlorocebus aethiops*) (Enkirch and von Messling, 2015; Finch et al., 2020; Gruber et al., 2021; Munster et al., 2020; Rockx et al., 2020). In these models there is no need for genetic modifications in order to study COVID-19 progression. Nevertheless, the models are less frequently used either due to lack of research tools, limited availability, high costs, complex husbandry or associated ethical concerns.

3.12 Murinised SARS-CoV2

While most efforts have been made in generating mouse models humanising the ACE2 to potentiate study of entry and infection *in vivo*, efforts have also been made in murinising the SARS-CoV2 virus itself. In a study by Muruato et al. (2021), they used a reverse genetic system and *in vivo* adaptation to successfully generate SARS-CoV2 strains capable of infecting mice (Muruato et al., 2021). Following infection of the murinised SARS-CoV2 strain the mouse lung exhibited substantial damage manifested with inflammation, immune infiltration, and pneumonia. The infection with the adapted virus was however only exhibited in the upper respiratory tract, thus is inappropriate for studies focusing on multisystem infection. It is worth mentioning that the novel adaptation of the virus was shown to keep its ability to

68

infect human airway cells (Muruato et al., 2021). This system, with a murinised SARS-CoV and a standard wild type laboratory mouse, overcomes tropism leading to encephalitis seen in infected transgenic mouse models whilst offering a system applicable for both *in vivo* mouse studies and *in vitro* studies on human primary cells (McCray et al., 2007; Winkler et al., 2020). In a study from 2020 the investigators had also produced a murinised SARS-CoV2 *via* reverse genetics to remodel the interaction between the mouse ACE2 and the virus which resulted in a recombinant virus able to infect the BALB/c mice. It was able replicate in both young and old mice however leading to more severe disease in older mice and exhibiting more clinically relevant phenotypes as compared to the disease presentation between non-modified SARS-CoV2 and transgenic mouse models (Dinnon et al., 2020). This gives the murinised ACE2 system better face validity, however the construct validity is decreased.

4 RECENT TRANSLATIONAL APPLICATIONS OF RODENT MODELS SUSCEPTIBLE TO SARS-COV2

The transgenic and transiently sensitised, humanised mouse models of SARS-CoV2 infection have gifted scientists the opportunity to study the potential destruction this, so far, relentless virus can cause to its host *in vivo*. Towards the beginning of the pandemic, initial studies using these models focused on the mechanisms in which viral entry can occur as well as their points of entry, the tissues primarily affected and the pathology of those tissues. These ongoing attempts to recreate infection have assisted our understanding of the infection timeline and has provided a guide to possible symptoms to be aware of in COVID-19 patients. Discussed here are animal studies performed in order to obtain risk assessments of new variants and evaluate the efficacy and safety of candidate anti-viral drugs for treatment in COVID-19 patients.

4.1 Risk Assessments of Variants of Concern

Particular mutations in the RBD have been key to identifying variants of concern. N501Y is one substitution that is characteristic of the Alpha (B.1.1.7) variant but is also found in the Beta (B.1.351), Gamma (B.1.1.28) and Omicron (B.1.1.529) variants (European Centre for Disease Prevention and Control, 2021; He et al., 2021). This means that N501Y is present in all but one variant of concern. *In silico* models predicted that this substitution occurs at a key residue for the RBD that is directly responsible for its strengthened affinity for ACE2 (Shahhosseini et al., 2021). The influence of N501Y was proved using a hamster model, where both donors and recipients inoculated with virus carrying N501Y showed significantly increased viral load in nasal washes and lung and trachea homogenates at 1–4 dpi compared to virus carrying a predecessor 'wild-type' spike protein. Substitutions S982A and D1118H were also shown to decrease viral fitness (Liu et al.,

2022). Interestingly, N501Y increases the infectivity of hosts expressing both either hACE2 or mACE2 (Pan et al., 2021), as it has been revealed that variants with this substitution possess an 8-fold higher affinity for the receptor (Bayarri-Olmos et al., 2021). While this demonstrates a key application of using animals in order to evaluate the potential potency of infection with rapidly evolving variants, we need more studies that apply these principles in the established transgenic and sensitised animal models expressing hACE2. This is because, ultimately, we will require data on how mutations in SARS-CoV2 will affect transmission between, and the health of, humans in the future.

Studies of this nature have been carried out. K18-hACE2 mice infected with B.1.1.7 show increased weight loss and hyperthermia earlier compared to mice exposed to B.1.351 or the initial WA-1 variant of concern. However, B.1.1.7 and B.1.351 infected mice display more severe clinical manifestations overall compared to WA-1 in a viral dose-dependent manner, with WA-1 infected mice displaying a 50% lower mortality rate at a dose of 10^3 pfu (Horspool et al., 2021). Again in K18-hACE2 mice, whilst both WA-1 and B.1.1.7 inoculated intranasally caused COVID-19-like disease in the mice, a lower dose of B.1.1.7 was required to cause a severe disease state (Bayarri-Olmos et al., 2021). In contrast, C57Bl/6J hACE2 knock-in mice display reduced viral load in lung and nasal turbinate, and a more minor lung pathology and inflammatory response on exposure to WA-1, B.1.1.7 or B.1.351 variants compared to the K18-hACE2 model, where viral RNA is concentrated at the epithelia of larger airways (Winkler et al., 2022). This is most likely attributed to the difference in approach of hACE2 expression in these two mouse lines and highlights the benefits of multiple rodent models of infection, chosen depending on the study focus, but also shows how different models could be affected when exposed to differing strains. Here, studies using models that more accurately recreate human infection to novel strains will possess increased extrapolative power.

The appearance of the B.1.1.529 (Omicron) variant in late 2021 came with heightened suspicions whether the current vaccines and therapies in progress would still provide suitable protection against a variant with >30 mutations in the RBD compared to variants described so far (Hodcroft, 2021). Halfmann et al. found that K18-hACE2 mice inoculated with B.1.351 showed significantly increased viral load in nasal turbinate and lung tissue homogenates at 3 dpi, and greater weight loss at 6 dpi compared to B.1.1.529 infected mice (Halfmann et al., 2022), suggesting reduced severity in viral manifestation on infection with the Omicron lineage in comparison to Beta lineage. These studies show the potential of exploiting the current animal models available in order to screen variants with specific mutations to assess their risk to humans, and for practitioners and governments to make appropriate decisions regarding patient care and infection control strategies.

4.2 Screening the Efficacy of Anti-Viral Therapies

The current pandemic has called for the development of anti-viral drugs in order to reduce or eliminate viral infection in, especially,

hospitalised COVID-19 patients. Due to the haste in which these drugs are required to ease the pressure of the pandemic on the world, drug development processes for SARS-CoV2 may be accelerated straight to clinical trials in humans, bypassing preclinical animal safety and efficacy studies.

Anti-viral molecules have been tested in a mouse setting though. PF-07304814 is a phosphate prodrug that on administration is processed into its active form PF-00835231, a potent cysteine protease inhibitor of coronavirus 3CL^{pro}, that was originally considered as a treatment for the 2002 SARS-CoV epidemic in 2003 (Hoffman et al., 2020). PF-00835231 is effective against alpha, beta, and gamma coronaviruses by preventing viral replication through inhibition of essential proteolysis by 3CL^{pro}. BALB/C mice infected with SARS-CoV2 MA10 display no weight loss and complete viral elimination when PF-00835231 is administered subcutaneously twice per day at a dose of 300 mg/kg. Initial weight loss is observed in mice receiving 30–100 mg/kg doses, which recovered to the starting weight at 4 dpi with viral load decreasing in dose-dependent manner. Significant decreases in viral load were also measured in SARS-COV2 exposed mice expressing hACE2 on treatment with PF-00835231. Additionally, this trend is also obtained even when treatment was delayed by 1dpi (Boras et al., 2021), highlighting the importance of identifying infection early, especially in high-risk patients. Despite hACE2 being expressed under a CMV promoter, which may not accurately follow the expected human expression of ACE2, this work shows the power of this inhibitor to prevent viral replication and poses a good option for further development into human clinics.

PF-07321332 (Nirmatrelvir) is another 3CL^{pro} inhibitor, which is the active component of the Pfizer-produced PAXLOVID™ (Pfizer, 2021), that gained approval in the UK (Medicines & Healthcare products Regulatory Agency, 2021a) and the United States (U.S. Food and Drug Administration, 2021a) at the end of 2021, and in the EU in January 2022 for treatment of COVID-19 (European Medicines Agency, 2022). An efficacy study in mice investigated the anti-viral activity of PF-07321332 in BALB/C mice infected with mouse-adapted SARS-CoV2 MA10. Mice treated via oral administration were protected from weight loss, had significantly reduced lung viral titre at 4 dpi and showed markedly decreased nucleocaspid presence in lung sections (Owen et al., 2021). Syrian hamsters have been shown to be protected from severe B.1.351 infection when treated with PF-07321332. Significant dose-dependent reductions in viral lung titre and improved weight retention at 4 dpi, as well as lung anatomy closely resembling uninfected hamsters was observed in those treated with PF-07321332. Hamsters were also completely protected from infection when co-housed for 2 days with a B.1.617.2 (Delta) variant positive cage mate when treated with PF-07321332 compared to those not (Abdelnabi et al., 2022). These rodent models support the continuing development and protective ability of PAXLOVID™ use in COVID-19 patients against multiple variants of concern, including the benefits of easy oral administration. Yet, further validation in humanised ACE2 rodent models, such as the hACE2 model used in Bao et al. (2020a) may be required for increased value in vivo, as the mentioned studies comprised of mouse adapted SARS-CoV2

infection and wild-type Syrian hamsters as part of their models. These studies could also be extended to examine potential side effects or long term ramifications for patients prescribed this anti-viral treatment. Synthesis and study of additional 3CL^{pro} inhibitors with favourable oral, intraperitoneal, and intravenous bioavailability have been reported and trialled in Sprague-Dawley rats and a CRISPR/Cas9 generated hACE2 expressing mice model (Qiao et al., 2021). However more work is required in this area, and PF-07321332 seems to have won the race for clinical trial approval.

Molnupiravir is another anti-viral drug that instead enforces a high mutagenesis rate via integration of its active form, β-D-N4-hydroxycytidine triphosphate (NHC), into viral RNA in the place of cytidine or uridine (Sheahan et al., 2020; Kabinger et al., 2021). It has so far gained approval for at-risk and hospitalised patients in the United States (U.S. Food and Drug Administration, 2021b) and the UK (Medicines & Healthcare products Regulatory Agency, 2021b). NHC shows potent viral inhibition and significantly reduces viral load in cell culture (Zhou et al., 2021) and diminishes weight loss, indicators of lung haemorrhage and lung viral titre at 500 mg/kg dosage in C57Bl/6 mice infected with either mouse adapted SARS-MA15 or MERS-CoV. Importantly, initiating NHC treatment before 24 h post infection showed to be crucial to maintaining reductions in weight loss, lung haemorrhaging, viral lung titre and lung and alveolar injury scores (Sheahan et al., 2020). However, a warning of mutagenic toxicity to host DNA during NHC treatment has been given, where mutations in a reporter gene increased in a dose-dependent manner with NHC. It has been suggested that the possible conversion of NHC to dNHC (2'-deoxyribose form of NHC) could be the cause of this increased mutational rate in the host genome (Zhou et al., 2021), and should be investigated further in an in vivo model focusing on tissues with natural proliferative tendency.

One study utilised immunodeficient mice with hACE2-and hTMPRSS2- expressing human lung tissue implanted in the animals' backs. This in vivo tissue model was susceptible to SARS-CoV, MERS-CoV and SARS-CoV2 infection, showed histopathological symptoms echoing viral damage and a 1000-fold increase in proinflammatory cytokines. Beginning NHC treatment at 12- and 24-h post-infection was extremely effective at reducing viral load, however if treatment started 12 h prior to infection, viral titre in the implanted human lung tissue was measured at >100,000-fold lower than the vehicle control, bestowing the protective potential of Molnupiravir in high-risk patients (Wahl et al., 2021). Molunipiravir-derived inhibitors have also shown to be effective at impeding SARS-CoV2 transmission in ferrets (Cox et al., 2021), and reducing viral replication and its associated lung pathologies in SARS-CoV2-susceptible Syrian hamsters, with amplified viral RNA mutations detected in hamsters that started treatment 12 h pre-infection compared to 12 h post-infection or vehicle control (Rosenke et al., 2021). When used in combination with Favipiravir, another anti-viral drug that acts through lethal mutagenesis but requires higher doses for optimal SARS-CoV2 suppression (Kaptein et al., 2020), hamsters treated with sub-optimal doses of a Molnupiravir/Favipiravir cocktail displayed lower viral loads

70

than hamsters treated with only one alone, with an implied additional transmission protection from cage-mates (Abdelnabi et al., 2021b).

These examples show that animal models can serve effectively in the screening of anti-virals in the current and future pandemics, and could assist in the recommendation of single or combinational therapies to complement human clinical trials. Finally, Molnupiravir has also shown its high protective ability against the B.1.1.7 and B.1.351 variants in Syrian hamsters (Abdelnabi et al., 2021a) and emphasises NHC's potent anti-viral mechanism is not dependent on specific sequences in the viral genome which may be mutated in future variants of concern, such is the case with e.g. monoclonal antibody treatment.

5 FUTURE RESEARCH DIRECTIONS

5.1 Risk Factors Suitable for Rodent Research

This pandemic has revealed that certain individuals are risk of developing severe COVID-19 illness or death. Factors such as age, male sex, and ethnicity have been attributed to a tendency to suffer from severe symptoms (Ebinger et al., 2020; Mughal et al., 2020; Williamson et al., 2020), as well as patients with comorbidities such as diabetes, hypertension and obesity (Alguwaihes et al., 2020; Ebinger et al., 2020; Huang et al., 2020; Li X. et al., 2020; Mughal et al., 2020; Williamson et al., 2020; Goyal et al., 2022). Whereas asthma may actually be protective (Avdeev et al., 2020; Skevaki et al., 2020; Zhu et al., 2020). Genetic or induced mouse models of these disease states are already well established, and the opportunity to combine transgenic and sensitised SARS-CoV2 models with models of human conditions potentially vulnerable to COVID-19 is waiting to be seized. This will allow us to further study comorbidities that may aggravate SARS-CoV2 transmission and pathophysiology, or contribute to any long-term damaging effects in humans. The human population is genetically and culturally diverse, but isolating comorbidities or genetic traits for study in a controlled environment will be vital.

5.1.1 Age

A report from early in the pandemic described that every additional 10 years of age associates with a 1.5-fold increased chance of requiring a higher level of hospitalised care during COVID-19 infection (Ebinger et al., 2020). SARS-COV2-related deaths peak in those aged 80+, who possess more than a 20-fold higher chance of death than those aged 50–59 (Williamson et al., 2020). This most likely attributed to an increase in comorbidities with age, even if yet to be detected. Rodents experience a much shorter life span than humans, making them an excellent model for studying age-related changes in COVID-19 research. ACE2 receptor expression has been described to both increase (Baker et al., 2021; Wark et al., 2021) and decrease (Chen et al., 2020; J.; Gu et al., 2021; Xudong et al., 2006; Yoon et al., 2016) with age in humans and rodents. However, Berni Canani et al. (2021) observed no significant differences in ACE2 expression between children <10 years old and adults 20–80 years old,

and Li M.-Y. et al. (2020) detected this same trend when comparing expression across multiple tissues in adults above or below 49 years of age. These contradicting reports suggest that ACE2 expression alone may not be a robust marker for identifying severe risk of SARS-CoV2 infection, and other factors in combination with ACE2 receptor expression must possess a decisive role. Comprehensive studies encompassing widespread tissue analysis of ACE2 expression in multiple age groups could well be accomplished to solve this, surely context-dependent, matter in rodent models of infection.

5.1.2 Diabetes

Diabetic patients are at increased risk of hospitalisation and mortality on infection with SARS-CoV2 (Ebinger et al., 2020), and are significantly more likely to require oxygen, intubation, antibiotics or dexamethasone on admission to hospital than non-diabetic patients (Alguwaihes et al., 2020). This is not entirely surprising considering increased cellular glucose levels assists in supporting viral replication (Codo et al., 2020). Overexpression of hACE2 boosts glucose tolerance and pancreatic β-cell function in diabetic mice (Bindom et al., 2010), whilst $ACE2^{-/y}$ knockout mice display impaired glucose tolerance alongside hepatic steatosis (Cao et al., 2016). Infection-induced downregulation of ACE2, and the resulting angiotensin II excess, therefore intensifies an already unbalanced glucose homeostasis. For these reasons, a bi-directional relationship between diabetes and COVID-19 infection has been proposed (Muniangi-Muhitu et al., 2020).

HFD-induced diabetic DPP4$^{H/M}$ male C57Bl/6 mice have been shown to be more vulnerable to severe signs of disease on infection with MERS-CoV when compared to lean controls, displaying prolonged weight loss and lung inflammation up to 21dpi (Kulcsar et al., 2019). Ma et al. (2021) however is the only study we found to date that has addressed the effect of the current SARS-CoV2 in a hACE2 expressing mouse model of diabetes. Ob/ob mice showed greater weight loss and increased lung immune infiltration when compared to non-diabetic mice at 5 dpi. Interestingly, this study also observed higher fasting blood glucose levels in both wild type and ob/ob mice infected with SARS-CoV2, compared to non-infected. Insulin tolerance was also non-significantly reduced in infected ob/ob mice (Ma et al., 2021). This is an alarming observation, and shows the potency of COVID-19 infection to disturb glucose homeostasis, not only in diabetic patients. Given that genetically-, chemically- or diet-induced rodent models of type 1 and type 2 diabetes are well established (King, 2012) in scientific literature, more research utilising hACE2-expressing rodents combined with these diabetic models will be extremely beneficial to understanding the risk posed on diabetic and non-diabetic people, both during and after contracting COVID-19.

5.1.3 Obesity

Obesity is another major risk factor for severe COVID-19 symptoms (Alguwaihes et al., 2020; Goyal et al., 2022) and COVID-19-related hospitalisation and mortality (Popkin et al., 2020). Its involvement in instigating this is likely intertwined with other comorbidities such as diabetes and hypertension. ACE2 is

71

expressed in subcutaneous and visceral adipose tissue (Al-Benna, 2020), and SARS-CoV2 nucleocapsids were detected in up to 5% of adipocytes in a small cohort of deceased COVID-19 patients (Basolo et al., 2022). Consequently, more adipose tissue will lead to surges in viral penetration and illness.

C57Bl/6 male mice fed a high-fat diet (HFD) display higher ACE2 expression in the lungs and trachea but reduced Tmprss2 expression in the oesophagus, whereas obese females display reduced ACE2 expression in the oesophagus and trachea with no differences in lung tissue (Sarver and Wong, 2021). This, at least in rodents, shows how obesity affects the expression of key SARS-CoV2 entry proteins differently in the two sexes, and should be investigated further in order to understand whether human patients should be treated according to sex. Further, HFD-fed rats see a 3.8- and 6-fold increase in lung Ace2 expression, and a 5.1- and 3.4-fold increase in Tmprss2 expression compared to standard and ketogenic diet fed rats, respectively. AT_1R and AT_2R levels were also significantly increased in HFD fed rats. Interestingly though, mice fed a ketogenic diet saw reduced AT_1R expression in pulmonary tissue compared to rats fed standard chow (da Eira et al., 2021), and this type of diet may help to safeguard diabetic or hypertensive humans. It would be meaningful to see further studies into the potential protective effects of certain diets on SARS-CoV2 infected rodent models.

The obesity-prone C57Bl/6N strain can provide valuable information in support of increased weight and diet on disease advancement and severity in SARS-CoV2 infected rodent models. Zhang et al., utilising leptin receptor dysfunctional C57BL/KsJ-db/db mice, observed a maintained 10% weight loss and more severe pulmonary pathology and inflammation in the obese model compared to db/+ controls inoculated with a mouse-adapted SARS-CoV2. Viral load was also significantly higher in obese lungs, nasal turbinates and trachea (Zhang et al., 2021). HFD-fed C57Bl/6N mice transduced with AdV-hACE2 also display more severe lung pathology than lean mice at 10 days post SARS-CoV2 infection, however a more comprehensive inflammatory profile should be included when studying models such as these (Rai et al., 2021).

This presented evidence further supports the role of obesity in severe COVID-19 patients, and in a way embodies the fusion of two pandemics. Researchers may now also look towards rodent models, preferably expressing hACE2 under its namesake promoter, to develop treatments to ease symptoms and reduce mortality in these patients in the short term. Patients may then turn to improve their diet and lifestyle habits post-recovery.

5.1.4 Hypertension

As the ACE2 receptor is responsible for initial SARS-CoV2 cell entry, it is logical that hypertension was among the top clinical presentations in patients suffering from severe COVID-19 (Huang et al., 2020; Li X. et al., 2020), through viral disruption of RAAS. ACE2 usually acts a negative regulator of RAAS, lowering blood pressure with anti-inflammatory effects. A number of RAAS modulators have been tested on rat primary *in vitro* cultures that principally act to increase or decrease ACE2 mRNA or protein levels (Hu et al., 2021), which in regards to a

COVID-19 patient may either encourage additional viral penetration or further increase blood pressure, respectively. ACE inhibitors or ARB drugs however, seem to display a protective effect against SARS-CoV2 infection and in-hospital mortality (Hippisley-Cox et al., 2020). Nevertheless, these types of studies mainly take into account hospital admissions and must account for a large number of comorbidities and variables.

Diet-induced obese C57Bl/6J mice display weight loss, improved glucose tolerance and reduced expression of inflammatory cytokines when treated with ACE inhibitors (Premaratna et al., 2012). $ACE^{-/-}$ mice show a similar trend (Jayasooriya et al., 2008), displaying the potential for this treatment to ease multiple COVID-19-related risk factors at once. There are a high number of inbred, outbred and transgenic rodent models used for hypertension research (Lerman et al., 2019). Jiang et al. (2022) recently published that SARS-CoV2 viral load in the lungs is higher in transgenic hACE2-expressing mice that have been induced into hypertension compared to normotensive hACE2 mice. Further, AT_1R blocker treatment improved lung pathology, reduced blood pressure and downregulated IL-6 and TNF-α expression in hypertensive hACE2 mice. This signifies that treatment provided protection to the organs on SARS-CoV2 infection overall, despite increased viral penetration in the heart and kidneys initially at early infection (Jiang et al., 2022).

A recent preprint article reported that the ACE inhibitor Lisinopril can raise the ACE2 expression landscape in the lungs, small intestine, kidney and brain of healthy mice, an effect that persists to at least 21 days post-termination of treatment (Brooks et al., 2022). Captopril, which also acts as an ACE inhibitor, appears to improve lung pathology and reduce inflammation during SARS-CoV2 infection in an angiotensin II-induced hypertensive and hACE2-expressing mouse model, without any detectable effect on viral load (Gao et al., 2021). These reports reinforce the potential but need for further clarification on RAAS modulators in COVID-19 research, but studies focusing specifically on hypertension in rodents on infection with the SARS-CoV2 virus are lacking. Nonetheless, with blood pressure measurements by techniques such as tail-cuff plethysmography and radiotelemetry readily available for use in rodents (Burger et al., 2014) and with a number of hypertension remedies on the market, future COVID-19 animal research focused on hypertension risk or the efficacy of RAAS modulators would benefit from integrating these methodologies into their study design for a greater *in vivo* view of hypertension in the current pandemic.

5.2 Insights Into Long-COVID

Post-acute COVID-19 sequelae or 'long-COVID' is a condition in which patients continue to suffer multiple COVID-19-related symptoms weeks or even months after testing negative for the virus, and can come in continuous or relapsing forms. The mechanisms behind symptom persistence are still unclear as presentation varies from patient to patient. Large scale studies from around the globe have witnessed exhaustive lists of symptoms (Davis et al., 2021; Hossain et al., 2021; C. Huang

72

et al., 2021; Pérez-González et al., 2022), with those who were hospitalised or required intensive care during primary infection especially at risk (Xie et al., 2022).

Rodent models of post-acute COVID-19 syndrome have been close to non-existent so far. This is likely due to subsided infections, or death, of animals in the models currently available, and the incorporation of early terminal analysis into experiment design. To more accurately study the long-term effects of SARS-CoV2 infection we require models that are even more 'comprehensively human' than those presented in **Table 1**, which more closely mimic aspects such as our own immune system. Researching viral infections by utilising non-human primates is an attractive option, due to marked similarities in physiology and immune responses to antigens with humans with the possibility for longitudinal studies in controlled environments (Estes et al., 2018). The rhesus macaque, African green monkey and pigtail macaque are susceptible to SARS-CoV2 infection and show mild-moderate COVID-19-associated lung pathologies (Clancy et al., 2021). Further, Böszörményi et al. (2021) observed that infected macaques show worsening lung lesions in CT scans, increases in specific cytokines in plasma, mild to moderate histopathological signs of pneumonia and the presence of viral RNA levels in a myriad of tissues up to 38 dpi, despite all subjects testing negative for SARS-CoV2 after 14 dpi. This suggests that these non-human primates are also susceptible to post-acute COVID-19 in a similar way to humans (Böszörményi et al., 2021).

For many researchers however, rodents are a preferred model based on their lower maintenance costs, shorter gestation period and the wealth of tools for transgenic manipulation. A promising example of a mouse model with a humanised immune system is MISTRG6. These immunodeficient mice express seven human cytokine genes knocked into their respective locus in the mouse's genome, and tolerate human hematopoietic stem cell engraftment (Rongvaux et al., 2014). MISTRG6 mice that transiently express hACE2 sustain prolonged viral titres and RNA, more severe lung pathology, and immune cell signatures to at least 35dpi of SARS-CoV2 compared to controls, emulating severe COVID-19 disease in humans. Convalescent plasma therapy showed a protective effect in these mice in regards to weight loss and viral clearance, however only prophylactic monoclonal antibody treatment improved prevention of T cell lung infiltration (Sefik et al., 2021). This again highlights the importance of early diagnosis in high risk patients, and it will be interesting to see more therapies tested on this model over longer time periods.

Finally, a recent preprint article has described their tracking of 10-weeks- and 1-year-old BALB/C mice for 120 days post infection with mouse adapted SARS-CoV2 MA10. Younger mice cleared infection twice as fast as older mice, with cytokine responses enduring in the latter until 30 dpi. Interestingly though, mice in the younger age group displayed a greater capability for tissue repair, and Molnupiravir was also effective at reducing disease prevalence in the older age group (Dinnon et al., 2022). Although this study is yet to be peer-reviewed, long-term mouse studies such as this will prove valuable in the fight against post-acute COVID-19 syndrome.

6 CONCLUDING REMARKS

The COVID-19 pandemic prompted a global scientific effort to produce a number of diverse animal models that mimic SARS-CoV2 infection. Mice have been and continue to be the preferred model organism used in scientific research due to their easy manipulation, short breeding time and genetic similarity to humans. However, the SARS-CoV2 virus binds inefficiently to the ACE2 receptor in mice thus preventing severe infection. This seemingly large barrier has been surmounted by the generation of transgenic and humanised mouse models. Additionally, efforts have been placed in reverse engineering the virus itself to increase its affinity to the mouse ACE2 receptor and causing COVID-19 symptoms. The aim of this review was to highlight individual COVID-19 mouse models and the tissue-specific replication of the virus and pathophysiology upon infection. We substantiate the review with examples of how these models have been used in regards to risk assessment of novel strains, developing therapeutics and elucidating the mechanisms of risk factors such as old age, diabetes, obesity and hypertension. We believe that this review can be used as a comprised guide for investigators researching which mouse model or which strategy to employ in regards to future COVID-19 research.

AUTHOR CONTRIBUTIONS

REFERENCES

Abdelnabi, R., Foo, C. S., de Jonghe, S., Maes, P., Weynand, B., and Neyts, J. (2021a). Molnupiravir Inhibits Replication of the Emerging SARS-CoV-2 Variants of Concern in a Hamster Infection Model. *J. Infect. Dis.* 224 (5), 749–753. doi:10.1093/infdis/jiab361

Abdelnabi, R., Foo, C. S., Jochmans, D., Vangeel, L., de Jonghe, S., Augustijns, P., et al. (2022). The Oral Protease Inhibitor (PF-07321332) Protects Syrian Hamsters against Infection with SARS-CoV-2 Variants of Concern. *Nat. Commun.* 13 (1). doi:10.1038/s41467-022-28354-0

Abdelnabi, R., Foo, C. S., Kaptein, S. J. F., Zhang, X., Do, T. N. D., Langendries, L., et al. (2021b). The Combined Treatment of Molnupiravir and Favipiravir

Results in a Potentiation of Antiviral Efficacy in a SARS-CoV-2 Hamster Infection Model. *EBioMedicine* 72, 103595. doi:10.1016/j.ebiom.2021.103595

Al-Benna, S. (2020). Association of High Level Gene Expression of ACE2 in Adipose Tissue with Mortality of COVID-19 Infection in Obese Patients. *Obes. Med.* 19, 100283. doi:10.1016/j.obmed.2020.100283

Alguwaihes, A. M., Al-Sofiani, M. E., Megdad, M., Albader, S. S., Alsari, M. H., Alelayan, A., et al. (2020). Diabetes and Covid-19 Among Hospitalized Patients in Saudi Arabia: a Single-centre Retrospective Study. *Cardiovasc. Diabetol.* 19 (1), 205. doi:10.1186/s12933-020-01184-4

Alsaadi, E. J. A., and Jones, I. M. (2019). Membrane Binding Proteins of Coronaviruses. *Future Virol.* 14 (4), 275–286. doi:10.2217/fvl-2018-0144

Asaka, M. N., Utsumi, D., Kamada, H., Nagata, S., Nakachi, Y., Yamaguchi, T., et al. (2021). Highly Susceptible SARS-CoV-2 Model in CAG Promoter-Driven

73

hACE2-Transgenic Mice. *JCI Insight* 6 (19), e152529. doi:10.1172/JCI.INSIGHT.152529

Avdeev, S., Moiseev, S., Brovko, M., Yavorovskiy, A., Umbetova, K., Akulkina, L., et al. (2020). Low Prevalence of Bronchial Asthma and Chronic Obstructive Lung Disease Among Intensive Care Unit Patients with COVID-19. *Allergy* 75 (10), 2703–2704. doi:10.1111/all.14420

Baker, S. A., Kwok, S., Berry, G. J., and Montine, T. J. (2021). Angiotensin-converting Enzyme 2 (ACE2) Expression Increases with Age in Patients Requiring Mechanical Ventilation. *PLoS ONE* 16 (2 February), e0247060. doi:10.1371/journal.pone.0247060

Bao, L., Deng, W., Huang, B., Gao, H., Liu, J., Ren, L., et al. (2020a). The Pathogenicity of SARS-CoV-2 in hACE2 Transgenic Mice. *Nature* 583 (7818), 830–833. doi:10.1038/s41586-020-2312-y

Bao, L., Gao, H., Deng, W., Lv, Q., Yu, H., Liu, M., et al. (2020b). Transmission of Severe Acute Respiratory Syndrome Coronavirus 2 via Close Contact and Respiratory Droplets Among Human Angiotensin-Converting Enzyme 2 Mice. *J. Infect. Dis.* 222 (4), 551–555. doi:10.1093/INFDIS/JIAA281

Basolo, A., Poma, A. M., Bonuccelli, D., Proietti, A., Macerola, E., Ugolini, C., et al. (2022). Adipose Tissue in COVID-19: Detection of SARS-CoV-2 in Adipocytes and Activation of the Interferon-Alpha Response. *J. Endocrinol. Invest.* 45, 1021–1029. doi:10.1007/s40618-022-01742-5

Bayarri-Olmos, R., Johnsen, L. B., Idorn, M., Reinert, L. S., Rosbjerg, A., Vang, S., et al. (2021). The alpha/B.1.1.7 SARS-CoV-2 Variant Exhibits Significantly Higher Affinity for ACE-2 and Requires Lower Inoculation Doses to Cause Disease in K18-hACE2 Mice. *ELife* 10, e70002. doi:10.7554/eLife.70002

Berni Canani, R., Comegna, M., Paparo, L., Cernera, G., Bruno, C., Strisciuglio, C., et al. (2021). Age-Related Differences in the Expression of Most Relevant Mediators of SARS-CoV-2 Infection in Human Respiratory and Gastrointestinal Tract. *Front. Pediatr.* 9. doi:10.3389/fped.2021.697390

Bindom, S. M., Hans, C. P., Xia, H., Boulares, A. H., and Lazartigues, E. (2010). Angiotensin I-Converting Enzyme Type 2 (ACE2) Gene Therapy Improves Glycemic Control in Diabetic Mice. *Diabetes* 59 (10), 2540–2548. doi:10.2337/db09-0782

Boras, B., Jones, R. M., Anson, B. J., Arenson, D., Aschenbrenner, L., Bakowski, M. A., et al. (2021). Preclinical Characterization of an Intravenous Coronavirus 3CL Protease Inhibitor for the Potential Treatment of COVID19. *Nat. Commun.* 12 (1), 6055. doi:10.1038/s41467-021-26239-2

Böszörményi, K. P., Stammes, M. A., Fagrouch, Z. C., Kiemenyi-Kayere, G., Niphuis, H., Mortier, D., et al. (2021). The post-acute Phase of Sars-Cov-2 Infection in Two Macaque Species Is Associated with Signs of Ongoing Virus Replication and Pathology in Pulmonary and Extrapulmonary Tissues. *Viruses* 13 (8), 1673. doi:10.3390/v13081673

Brooks, S. D., Smith, R. L., Moreira, A. S., and Ackerman, H. C. (2022). Oral Lisinopril Raises Tissue Levels of ACE2, The SARS-CoV-2 Receptor, in Healthy Male and Female Mice. *Front. Pharmacol.* 13, 798349. doi:10.3389/fphar.2022.798349

Bruter, A. V., Korshunova, D. S., Kubekina, M. V., Sergiev, P. V., Kalinina, A. A., Ilchuk, L. A., et al. (2021). Novel Transgenic Mice with Cre-dependent Co-expression of GFP and Human ACE2: a Safe Tool for Study of COVID-19 Pathogenesis. *Transgenic Res.* 30 (3), 289–301. doi:10.1007/S11248-021-00249-8/TABLES/1

Burger, D., Reudelhuber, T. L., Mahajan, A., Chibale, K., Sturrock, E. D., and Touyz, R. M. (2014). Effects of a Domain-Selective ACE Inhibitor in a Mouse Model of Chronic Angiotensin II-dependent Hypertension. *Clin. Sci.* 127 (1), 57–63. doi:10.1042/CS20130808

Calisher, C. H., Childs, J. E., Field, H. E., Holmes, K. V., and Schountz, T. (2006). Bats: Important Reservoir Hosts of Emerging Viruses. *Clin. Microbiol. Rev.* 19 (3), 531–545. doi:10.1128/CMR.00017-06

Cao, X., Yang, F., Shi, T., Yuan, M., Xin, Z., Xie, R., et al. (2016). Angiotensin-converting Enzyme 2/angiotensin-(1-7)/Mas axis Activates Akt Signaling to Ameliorate Hepatic Steatosis. *Sci. Rep.* 6, 21592. doi:10.1038/srep21592

Chan, J. F.-W., Yip, C. C.-Y., To, K. K.-W., Tang, T. H.-C., Wong, S. C.-Y., Leung, K.-H., et al. (2020). Improved Molecular Diagnosis of COVID-19 by the Novel, Highly Sensitive and Specific COVID-19-RdRp/Hel Real-Time Reverse Transcription-PCR Assay Validated *In Vitro* and with Clinical Specimens. *J. Clin. Microbiol.* 58 (5), e00310–20. doi:10.1128/JCM.00310-20

Chang, C.-k., Sue, S.-C., Yu, T.-h., Hsieh, C.-M., Tsai, C.-K., Chiang, Y.-C., et al. (2006). Modular Organization of SARS Coronavirus Nucleocapsid Protein. *J. Biomed. Sci.* 13 (1), 59–72. doi:10.1007/s11373-005-9035-9

Chen, J., Jiang, Q., Xia, X., Liu, K., Yu, Z., Tao, W., et al. (2020). Individual Variation of the SARS-CoV-2 Receptor ACE2 Gene Expression and Regulation. *Aging Cell* 19 (7), e13168. doi:10.1111/acel.13168

Clancy, C. S., Shaia, C., Munster, V., de Wit, E., Hawman, D., Okumura, A., et al. (2021). Histologic Pulmonary Lesions of SARS-CoV-2 in 4 Nonhuman Primate Species: An Institutional Comparative Review. *Vet. Pathol.*, 3009858211067468. doi:10.1177/03009858211067468

Codo, A. C., Davanzo, G. G., Monteiro, L. d. B., de Souza, G. F., Muraro, S. P., Virgilio-da-Silva, J. V., et al. (2020). Elevated Glucose Levels Favor SARS-CoV-2 Infection and Monocyte Response through a HIF-1α/Glycolysis-dependent Axis. *Cel Metab.* 32 (3), 437–446. e5. doi:10.1016/j.cmet.2020.07.007

Cox, R. M., Wolf, J. D., and Plemper, R. K. (2021). Therapeutically Administered Ribonucleoside Analogue MK-4482/EIDD-2801 Blocks SARS-CoV-2 Transmission in Ferrets. *Nat. Microbiol.* 6 (1), 11–18. doi:10.1038/s41564-020-00835-2

Czech Centre for Phenogenomics (2021). *Covid-19 Mouse Models Available at CCP - Czech Centre for Phenogenomics.* Vestec, Czechia. https://www.phenogenomics.cz/2021/09/covid-19-mouse-models-available-at-ccp/.

da Eira, D., Jani, S., and Ceddia, R. B. (2021). Obesogenic and Ketogenic Diets Distinctly Regulate the SARS-CoV-2 Entry Proteins ACE2 and TMPRSS2 and the Renin-Angiotensin System in Rat Lung and Heart Tissues. *Nutrients* 13 (10), 3357. doi:10.3390/nu13103357

Davis, H. E., Assaf, G. S., McCorkell, L., Wei, H., Low, R. J., Re'em, Y., et al. (2021). Characterizing Long COVID in an International Cohort: 7 Months of Symptoms and Their Impact. *EClinicalMedicine* 38, 101019. doi:10.1016/j.eclinm.2021.101019

De Gasparo, R., Pedotti, M., Simonelli, L., Nickl, P., Muecksch, F., Cassaniti, I., et al. (2021). Bispecific IgG Neutralizes SARS-CoV-2 Variants and Prevents Escape in Mice. *Nature* 593 (7859), 424–428. doi:10.1038/s41586-021-03461-y

Dinnon, K. H., III, Leist, S. R., Okuda, K., Dang, H., Fritch, E. J., Gully, K. L., et al. (2022). A Model of Persistent post SARS-CoV-2 Induced Lung Disease for Target Identification and Testing of Therapeutic Strategies. *BioRxiv.* doi:10.1101/2022.02.15.480515

Dinnon, K. H., Leist, S. R., Schäfer, A., Edwards, C. E., Martinez, D. R., Montgomery, S. A., et al. (2020). A Mouse-Adapted Model of SARS-CoV-2 to Test COVID-19 Countermeasures. *Nature* 586 (7830), 560–566. doi:10.1038/s41586-020-2708-8

Dolskiy, A. A., Gudymo, A. S., Taranov, O. S., Grishchenko, I. V., Shitik, E. M., Prokopov, D. Y., et al. (2022). The Tissue Distribution of SARS-CoV-2 in Transgenic Mice with Inducible Ubiquitous Expression of hACE2. *Front. Mol. Biosciences* 8, 1339. doi:10.3389/FMOLB.2021.821506/BIBTEX

Ebinger, J. E., Achamallah, N., Ji, H., Claggett, B. L., Sun, N., Botting, P., et al. (2020). Pre-existing Traits Associated with Covid-19 Illness Severity. *PLoS ONE* 15 (7 July), e0236240. doi:10.1371/journal.pone.0236240

Enkirch, T., and von Messling, V. (2015). Ferret Models of Viral Pathogenesis. *Virology* 479-480, 259–270. doi:10.1016/J.VIROL.2015.03.017

Estes, J. D., Wong, S. W., and Brenchley, J. M. (2018). Nonhuman Primate Models of Human Viral Infections. *Nat. Rev. Immunol.* 18 (6), 390–404. doi:10.1038/s41577-018-0005-7

European Centre for Disease Prevention and Control (2021). *Assessment of the Further Emergence of the SARS-CoV-2 Omicron VOC in the Context of the Ongoing Delta VOC Transmission in the EU/EEA, 18th Update.* Stockholm, Sweden: ECDC. Available at: https://www.ecdc.europa.eu/en/publications-data/covid-19-assessment-further-emergence-omicron-18th-risk-assessment (Accessed February 25, 2022).

European Medicines Agency (2021). *EMA Issues Advice on Use of Lagevrio (Molnupiravir) for the Treatment of COVID-19.* Amsterdam, Netherlands: European Medicines Agency. Available at: https://www.ema.europa.eu/en/news/ema-issues-advice-use-lagevrio-molnupiravir-treatment-covid-19 (Accessed February 10, 2022).

European Medicines Agency (2022). *Paxlovid.* European Medicines Agency. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/paxlovid (Accessed February 11, 2022).

74

Fehr, A. R., and Perlman, S. (2015). Coronaviruses: an Overview of Their Replication and Pathogenesis. *Methods Mol. Biol.* 1282, 1–23. doi:10.1007/978-1-4939-2438-7_1

Finch, C. L., Crozier, I., Lee, J. H., Byrum, R., Cooper, T. K., Liang, J., et al. (2020). Characteristic and Quantifiable COVID-19-like Abnormalities in CT- and PET/CT-imaged Lungs of SARS-CoV-2-Infected Crab-Eating Macaques (*Macaca fascicularis*). *BioRxiv.* doi:10.1101/2020.05.14.096727

Forni, D., Cagliani, R., Clerici, M., and Sironi, M. (2017). Molecular Evolution of Human Coronavirus Genomes. *Trends Microbiol.* 25 (1), 35–48. doi:10.1016/j.tim.2016.09.001

Gao, W.-C., Ma, X., Wang, P., Ma, X., Wang, P., He, X.-Y., et al. (2021). Captopril Alleviates Lung Inflammation in SARS-CoV-2-Infected Hypertensive Mice. *Zoolog. Res.* 42 (5), 633–636. doi:10.24272/J.ISSN.2095-8137.2021.206

Gavriatopoulou, M., Ntanasis-Stathopoulos, I., Korompoki, E., Fotiou, D., Migkou, M., Tzanninis, I.-G., et al. (2021). Emerging Treatment Strategies for COVID-19 Infection. *Clin. Exp. Med.* 21 (2), 167–179. doi:10.1007/s10238-020-00671-y

Goyal, A., Gupta, Y., Kalaivani, M., Praveen, P. A., Ambekar, S., and Tandon, N. (2022). SARS-CoV-2 Seroprevalence in Individuals with Type 1 and Type 2 Diabetes Compared with Controls. *Endocr. Pract.* 28 (2), 191–198. doi:10.1016/j.eprac.2021.12.009

Gruber, A. D., Firsching, T. C., Trimpert, J., and Dietert, K. (2021). Hamster Models of COVID-19 Pneumonia Reviewed: How Human Can They Be? *Vet. Pathol.*, 030098582110571. doi:10.1177/03009858211057197

Gu, H., Chen, Q., Yang, G., He, L., Fan, H., Deng, Y.-Q., et al. (2020). Rapid Adaptation of SARS-CoV-2 in BALB/c Mice: Novel Mouse Model for Vaccine Efficacy. *BioRxiv.* doi:10.1101/2020.05.02.073411

Gu, J., Yin, J., Zhang, M., Li, J., Wu, Y., Chen, J., et al. (2021). Study on the Clinical Significance of Ace2 and its Age-Related Expression. *Jir* Vol. 14, 2873–2882. doi:10.2147/JIR.S315981

Guo, T., Fan, Y., Chen, M., Wu, X., Zhang, L., He, T., et al. (2020). Cardiovascular Implications of Fatal Outcomes of Patients with Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol.* 5 (7), 811–818. doi:10.1001/jamacardio.2020.1017

Halfmann, P. J., Iida, S., Iwatsuki-Horimoto, K., Maemura, T., Kiso, M., Scheaffer, S. M., et al. (2022). SARS-CoV-2 Omicron Virus Causes Attenuated Disease in Mice and Hamsters. *Nature* 603, 687–692. doi:10.1038/s41586-022-04441-6

Hassan, A. O., Case, J. B., Winkler, E. S., Thackray, L. B., Kafai, N. M., Bailey, A. L., et al. (2020). A SARS-CoV-2 Infection Model in Mice Demonstrates Protection by Neutralizing Antibodies. *Cell* 182 (3), 744–753. doi:10.1016/J.CELL.2020.06.011

He, X., Hong, W., Pan, X., Lu, G., and Wei, X. (2021). SARS-CoV-2 Omicron Variant: Characteristics and Prevention. *MedComm* 2 (4), 838–845. doi:10.1002/mco2.110

Hippisley-Cox, J., Young, D., Coupland, C., Channon, K. M., Tan, P. S., Harrison, D. A., et al. (2020). Risk of Severe COVID-19 Disease with ACE Inhibitors and Angiotensin Receptor Blockers: Cohort Study Including 8.3 Million People. *Heart* 106 (19), 1503–1511. doi:10.1136/heartjnl-2020-317393

Hodcroft, E. B. (2021). *CoVariants: SARS-CoV-2 Mutations and Variants of Interest.* Available at https://covariants.org/ (Accessed February 10, 2022).

Hoffman, R. L., Kania, R. S., Brothers, M. A., Davies, J. F., Ferre, R. A., Gajiwala, K. S., et al. (2020). Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. *J. Med. Chem.* 63 (21), 12725–12747. doi:10.1021/acs.jmedchem.0c01063

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181 (2), 271–280. e8. doi:10.1016/j.cell.2020.02.052

Horspool, A. M., Ye, C., Wong, T. Y., Russ, B. P., Lee, K. S., Winters, M. T., et al. (2021). SARS-CoV-2 B.1.1.7 and B.1.351 Variants of Concern Induce Lethal Disease in K18-hACE2 Transgenic Mice Despite Convalescent Plasma Therapy. *BioRxiv.* doi:10.1101/2021.05.05.442784

Hossain, M. A., Hossain, K. M. A., Saunders, K., Uddin, Z., Walton, L. M., Raigangar, V., et al. (2021). Prevalence of Long COVID Symptoms in Bangladesh: A Prospective Inception Cohort Study of COVID-19 Survivors. *BMJ Glob. Health* 6 (12), e006838. doi:10.1136/bmjgh-2021-006838

Hu, Y., Liu, L., and Lu, X. (2021). Regulation of Angiotensin-Converting Enzyme 2: A Potential Target to Prevent COVID-19? *Front. Endocrinol.* 12. doi:10.3389/fendo.2021.725967

Huang, C., Huang, L., Wang, Y., Li, X., Ren, L., Gu, X., et al. (2021). 6-month Consequences of COVID-19 in Patients Discharged from Hospital: a Cohort Study. *The Lancet* 397 (10270), 220–232. doi:10.1016/S0140-6736(20)32656-8

Huang, S., Wang, J., Liu, F., Liu, J., Cao, G., Yang, C., et al. (2020). COVID-19 Patients with Hypertension Have More Severe Disease: a Multicenter Retrospective Observational Study. *Hypertens. Res.* 43 (8), 824–831. doi:10.1038/s41440-020-0485-2

Israelow, B., Song, E., Mao, T., Lu, P., Meir, A., Liu, F., et al. (2020). Mouse Model of SARS-CoV-2 Reveals Inflammatory Role of Type I Interferon Signaling. *J. Exp. Med.* 217 (12). doi:10.1084/JEM.20201241

Jackson, C. B., Farzan, M., Chen, B., and Choe, H. (2022). Mechanisms of SARS-CoV-2 Entry into Cells. *Nat. Rev. Mol. Cel Biol* 23 (1), 3–20. doi:10.1038/s41580-021-00418-x

Jayasooriya, A. P., Mathai, M. L., Walker, L. L., Begg, D. P., Denton, D. A., Cameron-Smith, D., et al. (2008). Mice Lacking Angiotensin-Converting Enzyme Have Increased Energy Expenditure, with Reduced Fat Mass and Improved Glucose Clearance. *Proc. Natl. Acad. Sci. U.S.A.* 105 (18), 6531–6536. www.pnas.orgcgi. doi:10.1073/pnas.0802690105

Jia, H., Yue, X., and Lazartigues, E. (2020). ACE2 Mouse Models: a Toolbox for Cardiovascular and Pulmonary Research. *Nat. Commun.* 11 (1), 5165. doi:10.1038/s41467-020-18880-0

Jiang, R.-D., Di, Liu, M.-Q., Liu, Y., Shan, C., Zhou, Y.-W., Shen, X.-R., et al. (2020). Pathogenesis of SARS-CoV-2 in Transgenic Mice Expressing Human Angiotensin-Converting Enzyme 2. *Cell* 182 (1), 50–58. e8. doi:10.1016/J.CELL.2020.05.027

Jiang, X., Li, H., Liu, Y., Bao, L., Zhan, L., Gao, H., et al. (2022). The Effects of ATIR Blocker on the Severity of COVID-19 in Hypertensive Inpatients and Virulence of SARS-CoV-2 in Hypertensive hACE2 Transgenic Mice. *J. Cardiovasc. Trans. Res.* 15, 38–48. doi:10.1007/s12265-021-10147-3

Kabinger, F., Stiller, C., Schmitzová, J., Dienemann, C., Kokic, G., Hillen, H. S., et al. (2021). Mechanism of Molnupiravir-Induced SARS-CoV-2 Mutagenesis. *Nat. Struct. Mol. Biol.* 28 (9), 740–746. doi:10.1038/s41594-021-00651-0

Kaptein, S. J. F., Jacobs, S., Langendries, L., Seldeslachts, L., ter Horst, S., Liesenborghs, L., et al. (2020). Favipiravir at High Doses Has Potent Antiviral Activity in SARS-CoV-2–infected Hamsters, whereas Hydroxychloroquine Lacks Activity. *Proc. Natl. Acad. Sci. U.S.A.* 117 (43), 26955–26965. doi:10.1073/pnas.2014441117/-/DCSupplemental

Katzman, C., Israely, T., Melamed, S., Politi, B., Sittner, A., Yahalom-Ronen, Y., et al. (2022). Modeling Sars-Cov-2 Infection in Mice Using Lentiviral Hace2 Vectors Infers Two Modes of Immune Responses to Sars-Cov-2 Infection. *Viruses* 14 (1), 11. doi:10.3390/V14010011/S1

King, A. J. (2012). The Use of Animal Models in Diabetes Research. *Br. J. Pharamacology* 166, 877–894. doi:10.1111/j.1476-5381.2012.01911.x

Kovacech, B., Fialova, L., Filipcik, P., Skrabana, R., Zilkova, M., Paulenka-Ivanovova, N., et al. (2022). Monoclonal Antibodies Targeting Two Immunodominant Epitopes on the Spike Protein Neutralize Emerging SARS-CoV-2 Variants of Concern. *EBioMedicine* 76, 103818. doi:10.1016/j10.1016/j.ebiom.2022.103818

Kulcsar, K. A., Coleman, C. M., Beck, S. E., and Frieman, M. B. (2019). Comorbid Diabetes Results in Immune Dysregulation and Enhanced Disease Severity Following MERS-CoV Infection. *JCI Insight* 4 (20), e131774. doi:10.1172/jci.insight.131774

Leclerc, Q. J., Fuller, N. M., Knight, L. E., Funk, S., Knight, G. M., and Knight, G. M. (2020). What Settings Have Been Linked to SARS-CoV-2 Transmission Clusters? *Wellcome Open Res.* 5, 83. doi:10.12688/wellcomeopenres.15889.1

Lerman, L. O., Kurtz, T. W., Touyz, R. M., Ellison, D. H., Chade, A. R., Crowley, S. D., et al. (2019). Animal Models of Hypertension: A Scientific Statement from the American Heart Association. *Hypertension* 73 (6), e87–e120. doi:10.1161/HYP.0000000000000090

Li, M.-Y., Li, L., Zhang, Y., and Wang, X.-S. (2020). Expression of the SARS-CoV-2 Cell Receptor Gene ACE2 in a Wide Variety of Human Tissues. *Infect. Dis. Poverty* 9 (1). doi:10.1186/s40249-020-00662-x

75

Li, X., Xu, S., Yu, M., Wang, K., Tao, Y., Zhou, Y., et al. (2020). Risk Factors for Severity and Mortality in Adult COVID-19 Inpatients in Wuhan. *J. Allergy Clin. Immunol.* 146 (1), 110–118. doi:10.1016/j.jaci.2020.04.006

Liu, Y., Liu, J., Plante, K. S., Plante, J. A., Xie, X., Zhang, X., et al. (2022). The N501Y Spike Substitution Enhances SARS-CoV-2 Infection and Transmission. *Nature* 602 (7896), 294–299. doi:10.1038/s41586-021-04245-0

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., et al. (2020). Genomic Characterisation and Epidemiology of 2019 Novel Coronavirus: Implications for Virus Origins and Receptor Binding. *The Lancet* 395 (10224), 565–574. doi:10.1016/s0140-6736(20)30251-8

Ma, Y., Lu, D., Bao, L., Qu, Y., Liu, J., Qi, X., et al. (2021). SARS-CoV-2 Infection Aggravates Chronic Comorbidities of Cardiovascular Diseases and Diabetes in Mice. *Anim. Models Exp. Med.* 4 (1), 2–15. doi:10.1002/ame2.12155

McCray, P. B., Pewe, L., Wohlford-Lenane, C., Hickey, M., Manzel, L., Shi, L., et al. (2007). Lethal Infection of K18- hACE2 Mice Infected with Severe Acute Respiratory Syndrome Coronavirus. *J. Virol.* 81 (2), 813–821. doi:10.1128/JVI.02012-06

Medicines & Healthcare products Regulatory Agency (2021b). *First Oral Antiviral for COVID-19, Lagevrio (Molnupiravir), Approved by MHRA. [Press release]* London, United Kingdom. Available at: https://www.gov.uk/government/news/first-oral-antiviral-for-covid-19-lagevrio-molnupiravir-approved-by-mhra (Accessed February 24, 2022).

Medicines & Healthcare products Regulatory Agency (2021a). *Oral COVID-19 Antiviral, Paxlovid, Approved by UK regulator. [Press release].* London, United Kingdom (Available at: https://www.gov.uk/government/news/oral-covid-19-antiviral-paxlovid-approved-by-uk-regulator (Accessed February 24, 2022).

Menachery, V. D., Yount, B. L., Sims, A. C., Debbink, K., Agnihothram, S. S., Gralinski, L. E., et al. (2016). SARS-like WIV1-CoV Poised for Human Emergence. *Proc. Natl. Acad. Sci. U.S.A.* 113 (11), 3048–3053. doi:10.1073/PNAS.1517719113

Mughal, M. S., Kaur, I. P., Jaffery, A. R., Dalmacion, D. L., Wang, C., Koyoda, S., et al. (2020). COVID-19 Patients in a Tertiary US Hospital: Assessment of Clinical Course and Predictors of the Disease Severity. *Respir. Med.* 172, 106130. doi:10.1016/j.rmed.2020.106130

Muniangi-Muhitu, H., Akalestou, E., Salem, V., Misra, S., Oliver, N. S., and Rutter, G. A. (2020). Covid-19 and Diabetes: A Complex Bidirectional Relationship. *Front. Endocrinol.* 11. doi:10.3389/fendo.2020.582936

Munster, V. J., Feldmann, F., Williamson, B. N., van Doremalen, N., Pérez-Pérez, L., Schulz, J., et al. (2020). Respiratory Disease in Rhesus Macaques Inoculated with SARS-CoV-2. *Nature* 585 (7824), 268–272. doi:10.1038/S41586-020-2324-7

Muruato, A., Vu, M. N., Johnson, B. A., Davis-Gardner, M. E., Vanderheiden, A., Lokugamage, K., et al. (2021). Mouse-adapted SARS-CoV-2 Protects Animals from Lethal SARS-CoV challenge. *Plos Biol.* 19 (11), e3001284. doi:10.1371/journal.pbio.3001284

Nehme, Z., Zouein, F. A., Zayeri, Z. D., and Zibara, K. (2019). An Update on the Tissue Renin Angiotensin System and its Role in Physiology and Pathology. *Jcdd* 6 (2), 14. doi:10.3390/jcdd6020014

Owen, D. R., Allerton, C. M. N., Anderson, A. S., Aschenbrenner, L., Avery, M., Berritt, S., et al. (2021). An Oral SARS-CoV-2 M Pro Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science* 374 (6575), 1586–1593. https://www.science.org. doi:10.1126/science.abl4784

Pan, T., Chen, R., He, X., Yuan, Y., Deng, X., Li, R., et al. (2021). Infection of Wild-type Mice by SARS-CoV-2 B.1.351 Variant Indicates a Possible Novel Cross-Species Transmission Route. *Sig Transduct Target. Ther.* 6 (1), 420. doi:10.1038/s41392-021-00848-1

Pérez-González, A., Araújo-Ameijeiras, A., Fernández-Villar, A., Crespo, M., Poveda, E., Cabrera, J. J., et al. (2022). Long COVID in Hospitalized and Non-hospitalized Patients in a Large Cohort in Northwest Spain, a Prospective Cohort Study. *Sci. Rep.* 12 (1), 3369. doi:10.1038/s41598-022-07414-x

Popkin, B. M., Du, S., Green, W. D., Beck, M. A., Algaith, T., Herbst, C. H., et al. (2020). Individuals with Obesity and COVID-19: A Global Perspective on the Epidemiology and Biological Relationships. *Obes. Rev.* 21 (11), e13128. doi:10.1111/obr.13128

Premaratna, S. D., Manickam, E., Begg, D. P., Rayment, D. J., Hafandi, A., Jois, M., et al. (2012). Angiotensin-converting Enzyme Inhibition Reverses Diet-Induced Obesity, Insulin Resistance and Inflammation in C57BL/6J Mice. *Int. J. Obes.* 36 (2), 233–243. doi:10.1038/ijo.2011.95

Qiao, J., Li, Y.-S., Zeng, R., Liu, F.-L., Luo, R.-H., Huang, C., et al. (2021). SARS-CoV-2 M Pro Inhibitors with Antiviral Activity in a Transgenic Mouse Model. *Science* 371 (6536), 1374–1378. https://pymol.org. doi:10.1126/science.abf1611

Qing, E., and Gallagher, T. (2020). SARS Coronavirus Redux. *Trends Immunol.* 41 (4), 271–273. doi:10.1016/j.it.2020.02.007

Ragab, D., Salah Eldin, H., Taeimah, M., Khattab, R., and Salem, R. (2020). The COVID-19 Cytokine Storm; what We Know So Far. *Front. Immunol.* 11. doi:10.3389/fimmu.2020.01446

Rai, P., Chuong, C., LeRoith, T., Smyth, J. W., Panov, J., Levi, M., et al. (2021). Adenovirus Transduction to Express Human ACE2 Causes Obesity-specific Morbidity in Mice, Impeding Studies on the Effect of Host Nutritional Status on SARS-CoV-2 Pathogenesis. *Virology* 563, 98–106. doi:10.1016/j.virol.2021.08.014

Rathnasinghe, R., Strohmeier, S., Amanat, F., Gillespie, V. L., Krammer, F., García-Sastre, A., et al. (2020). Comparison of Transgenic and Adenovirus hACE2 Mouse Models for SARS-CoV-2 Infection. *BioRxiv.* doi:10.1101/2020.07.06.190066

Rawle, D. J., Le, T. T., Dumenil, T., Yan, K., Tang, B., Nguyen, W., et al. (2021). ACE2-lentiviral Transduction Enables Mouse SARS-CoV-2 Infection and Mapping of Receptor Interactions. *Plos Pathog.* 17 (7), e1009723. doi:10.1371/JOURNAL.PPAT.1009723

Rockx, B., Kuiken, T., Herfst, S., Bestebroer, T., Lamers, M. M., Oude Munnink, B. B., et al. (2020). Comparative Pathogenesis of COVID-19, MERS, and SARS in a Nonhuman Primate Model. *Science* 368 (6494), 1012–1015. doi:10.1126/SCIENCE.ABB7314

Rongvaux, A., Willinger, T., Martinek, J., Strowig, T., Gearty, S. v., Teichmann, L. L., et al. (2014). Development and Function of Human Innate Immune Cells in a Humanized Mouse Model. *Nat. Biotechnol.* 32 (4), 364–372. doi:10.1038/nbt.2858

Rosenke, K., Hansen, F., Schwarz, B., Feldmann, F., Haddock, E., Rosenke, R., et al. (2021). Orally Delivered MK-4482 Inhibits SARS-CoV-2 Replication in the Syrian Hamster Model. *Nat. Commun.* 12 (1). doi:10.1038/s41467-021-22580-8

Samavati, L., and Uhal, B. D. (2020). ACE2, Much More Than Just a Receptor for SARS-COV-2. *Front. Cell Infect. Microbiol.* 10. doi:10.3389/fcimb.2020.00317

Sarver, D. C., and Wong, G. W. (2021). Obesity Alters Ace2 and Tmprss2 Expression in Lung, Trachea, and Esophagus in a Sex-dependent Manner: Implications for COVID-19. *Biochem. Biophysical Res. Commun.* 538, 92–96. doi:10.1016/j.bbrc.2020.10.066

Sefik, E., Israelow, B., Mirza, H., Zhao, J., Qu, R., Kaffe, E., et al. (2021). A Humanized Mouse Model of Chronic COVID-19. *Nat. Biotechnol.* doi:10.1038/s41587-021-01155-4

Sellers, R. S. (2017). Translating Mouse Models. *Toxicol. Pathol.* 45 (1), 134–145. doi:10.1177/0192623316675767

Shahhosseini, N., Babuadze, G., Wong, G., and Kobinger, G. P. (2021). Mutation Signatures and In Silico Docking of Novel Sars-Cov-2 Variants of Concern. *Microorganisms* 9 (5), 926. doi:10.3390/microorganisms9050926

Sheahan, T. P., Sims, A. C., Zhou, S., Graham, R. L., Pruijssers, A. J., Agostini, M. L., et al. (2020). An Orally Bioavailable Broad-Spectrum Antiviral Inhibits SARS-CoV-2 in Human Airway Epithelial Cell Cultures and Multiple Coronaviruses in Mice. *Sci. Transl. Med.* 12 (541). doi:10.1126/SCITRANSLMED.ABB5883

Shou, S., Liu, M., Yang, Y., Kang, N., Song, Y., Tan, D., et al. (2021). Animal Models for COVID-19: Hamsters, Mouse, Ferret, Mink, Tree Shrew, and Non-human Primates. *Front. Microbiol.* 12. doi:10.3389/fmicb.2021.626553

Shuai, L., Zhong, G., Yuan, Q., Wen, Z., Wang, C., He, X., et al. (2021). Replication, Pathogenicity, and Transmission of SARS-CoV-2 in Minks. *Natl. Sci. Rev.* 8 (3), 2021. doi:10.1093/NSR/NWAA291

Simmons, G., Gosalia, D. N., Rennekamp, A. J., Reeves, J. D., Diamond, S. L., and Bates, P. (2005). Inhibitors of Cathepsin L Prevent Severe Acute Respiratory Syndrome Coronavirus Entry. *Proc. Natl. Acad. Sci. U.S.A.* 102 (33), 11876–11881. doi:10.1073/pnas.0505577102

Skevaki, C., Karsonova, A., Karaulov, A., Xie, M., and Renz, H. (2020). Asthma-associated Risk for COVID-19 Development. *J. Allergy Clin. Immunol.* 146 (6), 1295–1301. doi:10.1016/j.jaci.2020.09.017

Su, Y., Yuan, D., Chen, D. G., Ng, R. H., Wang, K., Choi, J., et al. (2022). Multiple Early Factors Anticipate post-acute COVID-19 Sequelae. *Cell* 185, 881–895. doi:10.1016/j.cell.2022.01.014

76

Sun, J., Zhuang, Z., Zheng, J., Li, K., Wong, R. L.-Y., Liu, D., et al. (2020). Generation of a Broadly Useful Model for COVID-19 Pathogenesis, Vaccination, and Treatment. *Cell* 182 (3), 734–743. e5. doi:10.1016/j.cell.2020.06.010

Sun, S.-H., Chen, Q., Gu, H.-J., Yang, G., Wang, Y.-X., Huang, X.-Y., et al. (2020). A Mouse Model of SARS-CoV-2 Infection and Pathogenesis. *Cell Host and Microbe* 28 (1), 124–133. e4. doi:10.1016/J.CHOM.2020.05.020/ATTACHMENT/34A485A8-CAD9-409E-AA5B-0EF7B1B54781/MMC1.PDF

Tang, C., Deng, Z., Li, X., Yang, M., Tian, Z., Chen, Z., et al. (2020). Helicase of Type 2 Porcine Reproductive and Respiratory Syndrome Virus Strain HV Reveals a Unique Structure. *Viruses* 12 (2), 215. doi:10.3390/v12020215

Thomas, S. (2020). The Structure of the Membrane Protein of SARS-CoV-2 Resembles the Sugar Transporter SemiSWEET. *Pai* 5 (1), 342–363. doi:10.20411/pai.v5i1.377

Triggle, C. R., Bansal, D., Ding, H., Islam, M. M., Farag, E. A. B. A., Hadi, H. A., et al. (2021). A Comprehensive Review of Viral Characteristics, Transmission, Pathophysiology, Immune Response, and Management of SARS-CoV-2 and COVID-19 as a Basis for Controlling the Pandemic. *Front. Immunol.* 12, 631139. doi:10.3389/fimmu.2021.631139

Tseng, C.-T. K., Huang, C., Newman, P., Wang, N., Narayanan, K., Watts, D. M., et al. (2007). Severe Acute Respiratory Syndrome Coronavirus Infection of Mice Transgenic for the Human Angiotensin-Converting Enzyme 2 Virus Receptor. *J. Virol.* 81 (3), 1162–1173. doi:10.1128/JVI.01702-06/ASSET/DACA1C30-425C-4D1C-8B34-1CE2B664B6C9/ASSETS/GRAPHIC/ZJV0030787060008.JPEG

Turner, A. J., and Hooper, N. M. (2002). The Angiotensin-Converting Enzyme Gene Family: Genomics and Pharmacology. *Trends Pharmacol. Sci.* 23 (4), 177–183. doi:10.1016/s0165-6147(00)01994-5

U.S. Food and Drug Administration (2021b). *Coronavirus (COVID-19) Update: FDA Authorizes Additional Oral Antiviral for Treatment of COVID-19 in Certain Adults.* Silver Spring, MD: FDA. Available at: https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-additional-oral-antiviral-treatment-covid-19-certain (Accessed February 11, 2022).

U.S. Food and Drug Administration (2021a). *Coronavirus (COVID-19) Update: FDA Authorizes First Oral Antiviral for Treatment of COVID-19.* Silver Spring, MD: FDA. Available at: https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-oral-antiviral-treatment-covid-19 (Accessed February 10, 2022).

Wahl, A., Gralinski, L. E., Johnson, C. E., Yao, W., Kovarova, M., Dinnon, K. H., et al. (2021). SARS-CoV-2 Infection Is Effectively Treated and Prevented by EIDD-2801. *Nature* 591 (7850), 451–457. doi:10.1038/s41586-021-03312-w

Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181 (2), 281–292. e6. doi:10.1016/j.cell.2020.02.058

Wark, P. A. B., Pathinayake, P. S., Kaiko, G., Nichol, K., Ali, A., Chen, L., et al. (2021). ACE2 Expression Is Elevated in Airway Epithelial Cells from Older and Male Healthy Individuals but Reduced in Asthma. *Respirology* 26 (5), 442–451. doi:10.1111/resp.14003

Williamson, E. J., Walker, A. J., Bhaskaran, K., Bacon, S., Bates, C., Morton, C. E., et al. (2020). Factors Associated with COVID-19-Related Death Using OpenSAFELY. *Nature* 584 (7821), 430–436. doi:10.1038/s41586-020-2521-4

Winkler, E. S., Chen, R. E., Alam, F., Yildiz, S., Case, J. B., Uccellini, M. B., et al. (2022). SARS-CoV-2 Causes Lung Infection without Severe Disease in Human ACE2 Knock-In Mice. *J. Virol.* 96 (1), e0151121. doi:10.1128/JVI.01511-21

Winkler, E. S., Bailey, A. L., Kafai, N. M., Nair, S., McCune, B. T., Yu, J., et al. (2020). SARS-CoV-2 Infection of Human ACE2-Transgenic Mice Causes Severe Lung Inflammation and Impaired Function. *Nat. Immunol.* 21 (11), 1327–1335. doi:10.1038/s41590-020-0778-2

Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C.-L., Abiona, O., et al. (2020). Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation. *Science* 367 (6483), 1260–1263. doi:10.1126/science.abb2507

Xie, Y., Xu, E., Bowe, B., and Al-Aly, Z. (2022). Long-term Cardiovascular Outcomes of COVID-19. *Nat. Med.* 28, 583–590. doi:10.1038/s41591-022-01689-3

Xudong, X., Junzhu, C., Xingxiang, W., Furong, Z., and Yanrong, L. (2006). Age- and Gender-Related Difference of ACE2 Expression in Rat Lung. *Life Sci.* 78 (19), 2166–2171. doi:10.1016/j.lfs.2005.09.038

Yang, X. H., Deng, W., Tong, Z., Liu, Y. X., Zhang, L. F., Zhu, H., et al. (2007b). Mice Transgenic for Human Angiotensin-Converting Enzyme 2 Provide a Model for SARS Coronavirus Infection. *Comp. Med.* 57 (5), 450–459.

Yang, X. H., Deng, W., Tong, Z., Liu, Y.-x., Zhang, L., and Zhu, H. (2007a). Mice Transgenic for Human Angiotensin-Converting Enzyme 2 Provide a Model for SARS Coronavirus Infection. *Comp. Med.* 57, 450–459. *Ingenta Connect.* from https://www.ingentaconnect.com/content/aalas/cm/2007/00000057/00000005/art00003# (Retrieved March 6, 2022).

Yong, S. J. (2021). Long COVID or post-COVID-19 Syndrome: Putative Pathophysiology, Risk Factors, and Treatments. *Infect. Dis.* 53 (Issue 10), 737–754. doi:10.3744235.2021.1924397

Yoon, H. E., Kim, E. N., Kim, M. Y., Lim, J. H., Jang, I.-A., Ban, T. H., et al. (2016). Age-Associated Changes in the Vascular Renin-Angiotensin System in Mice. *Oxidative Med. Cell Longevity* 2016, 1–14. doi:10.1155/2016/6731093

Zhang, X., Kousoulas, K. G., and Storz, J. (1992). The Hemagglutinin/esterase Gene of Human Coronavirus Strain OC43: Phylogenetic Relationships to Bovine and Murine Coronaviruses and Influenza C Virus. *Virology* 186 (1), 318–323. doi:10.1016/0042-6822(92)90089-8

Zhang, Y.-N., Zhang, Z.-R., Zhang, H.-Q., Li, X.-D., Li, J.-Q., Zhang, Q.-Y., et al. (2021). Increased Morbidity of Obese Mice Infected with Mouse-Adapted SARS-CoV-2. *Cell Discov* 7 (Issue 1), 74. doi:10.1038/s41421-021-00305-x

Zhou, S., Hill, C. S., Sarkar, S., Tse, L. v., Woodburn, B. M. D., Schinazi, R. F., et al. (2021). β-d-N4-hydroxycytidine Inhibits SARS-CoV-2 through Lethal Mutagenesis but Is Also Mutagenic to Mammalian Cells. *J. Infect. Dis.* 224 (3), 415–419. doi:10.1093/infdis/jiab247

Zhu, Z., Hasegawa, K., Ma, B., Fujiogi, M., Camargo, C. A., Jr., and Liang, L. (2020). Association of Asthma and its Genetic Predisposition with the Risk of Severe COVID-19. *J. Allergy Clin. Immunol.* 146 (2), 327–e4. doi:10.1016/j.jaci.2020.05.02110.1016/j.jaci.2020.06.001

77

4.3 Part 2. Application of rAAV system in disease modelling

Site-specific recombinase (SSR) systems, including Cre/*loxP*, Flpo/*frt*, Dre/*rox*, and Vika/*vox*, provide precise control over tissue-specific and/or time-dependent recombination events in biological systems. Among these, the Cre/*loxP* system is the most widely used for studying complex biological processes via genetic manipulation. Building on our previous work in efficient and direct allele conversion with Cre protein, as demonstrated by Jenickova *et al.* (2021), we sought to expand our SSR protein portfolio. While we successfully synthesized the Dre protein, attempts to produce Flp and Vika recombinases in sufficient quantities were unsuccessful. To overcome this limitation, we packaged the SSRs into recombinant AAV (rAAV) vectors.

These rAAV vectors were used to treat zygotes and embryos *ex vivo*, facilitating the delivery and expression of all SSRs in mouse embryos with high conversion efficiency. This method also allowed for multi-level allele conversion using two distinct recombinases. Moreover, the low persistence of AAV genomes enabled the use of F0 animals for further experimental work, significantly reducing the number of animals required and shortening the time needed to obtain converted allele-positive and SSR-free animals. In summary, this article presents a novel method for allele conversion using rAAV vectors and compares it to conventional techniques, such as electroporation of SSR proteins or mRNA microinjection.

scientific reports

OPEN

Multistep allelic conversion in mouse pre-implantation embryos by AAV vectors

Petr Nickl[1,2✉], Irena Jenickova[1], Jan Elias[1,2], Petr Kasparek[1,2], Cyril Barinka[3], Jana Kopkanova[1] & Radislav Sedlacek[1,2✉]

Site-specific recombinases (SSRs) are critical for achieving precise spatiotemporal control of engineered alleles. These enzymes play a key role in facilitating the deletion or inversion of loci flanked by recombination sites, resulting in the activation or repression of endogenous genes, selection markers or reporter elements. However, multiple recombination in complex alleles can be laborious. To address this, a new and efficient method using AAV vectors has been developed to simplify the conversion of systems based on Cre, FLP, Dre and Vika recombinases. In this study, we present an effective method for ex vivo allele conversion using Cre, FLP (flippase), Dre, and Vika recombinases, employing adeno-associated viruses (AAV) as delivery vectors. AAVs enable efficient allele conversion with minimal toxicity in a reporter mouse line. Moreover, AAVs facilitate sequential allele conversion, essential for fully converting alleles with multiple recombination sites, typically found in conditional knockout mouse models. While simple allele conversions show a 100% efficiency rate, complex multiple conversions consistently achieve an 80% conversion rate. Overall, this strategy markedly reduces the need for animals and significantly speeds up the process of allele conversion, representing a significant improvement in genome engineering techniques.

Site-specific recombinase (SSRs) systems represent powerful tools in genetic engineering and molecular biology, enabling precise spatiotemporal manipulation of DNA sequences. The two most widely employed systems are the Cre/*loxP* system, where Cre recombinase binds *loxP* sites, and the FLP/*FRT* system, which involves Flpo recombinase (a mammalian codon-optimized thermostable version of the FLP gene) in the conversion of FRT sites in mammalian genomes. In addition, alternative systems such as Dre/*rox* and Vika/*vox* are part of this SSR toolbox. All these systems facilitate site-specific recombination, allowing for the excision, inversion, insertion, or exchange of DNA segments[1].

The Cre/*loxP* system is the major genetic tool used for various genetic manipulations, including gene knockout, conditional gene inactivation, and transgenic cassette rearrangements[2,3]. The Cre/*loxP* system, in combination with FLP/*FRT*, Dre/*rox* or Vika/*vox,* allows for complex genetic manipulation[4]. The inclusion of Dre and Vika systems has further enriched the toolbox for combinatorial use, offering enhanced flexibility and precision in genetic manipulations[5]. Despite its limited application, the Vika/*vox* system has demonstrated suitability as an alternative or a combinatorial partner with other site-specific recombinase (SSR) systems in vivo, as demonstrated by Karimova et al.[6,7]. The Cre/*loxP* and Dre/*rox* systems are widely combined for various purposes, as shown, for example, in vivo lineage tracing[8,9]. Similarly, the combination of Cre/*loxP* and FLP/*FRT* systems serves as the foundation for gene manipulation using the EUCOMM/KOMP knockout first alleles[3].

The EUCOMM/KOMP "knockout-first" allele utilizes strategic placing of recognition sequences in the vicinity of critical exons and gives the gene of interest conditional properties. This enables functional ablation of a target gene in a spatiotemporal manner. EUCOMM/KOMP alleles find extensive use in generating genetically modified rodents with specific gene reporters, knockouts, or conditional knockouts based on the applied SSR[3,10]. EUCOMM/KOMP employs promoter-driven targeting selection cassettes to create a versatile 'Knockout-first allele' in C57BL/6N embryonic stem cells[3,10]. This innovative strategy hinges on identifying a 'critical' exon

[1]Czech Centre for Phenogenomics, Institute of Molecular Genetics of the Czech Academy of Sciences, Prumyslova 595, 252 50 Vestec, Czech Republic. [2]Laboratory of Transgenic Models of Diseases, Institute of Molecular Genetics of the Czech Academy of Sciences, Videnska 1083, 142 20 Prague, Czech Republic. [3]Institute of Biotechnology of the Czech Academy of Sciences, BIOCEV, Prumyslova 595, 252 50 Vestec, Czech Republic. ✉email: petr.nickl@img.cas.cz; radislav.sedlacek@img.cas.cz

1

present in all transcript variants. In its final form (tm1b or tm1d), the deletion of the exon induces a frame-shift mutation and thus forms the KO-first allele. Notably, the allele exhibits flexibility, allowing for the generation of reporter knockouts (tm1a), conditional knockouts (tm1c), and null alleles through exposure to site-specific recombinases such as Cre and FLP.

The 'knockout-first' allele (tm1a) contains an *IRES:lacZ* trapping cassette and a floxed promoter-driven neo cassette inserted into the intron of a gene, disrupting gene function. FLP converts the 'knockout-first' allele to a conditional allele (tm1c), restoring gene activity. Cre deletes the promoter-driven selection cassette and floxed exon of the tm1a allele to generate a *lacZ*-tagged allele (tm1b) or deletes the floxed exon of the tm1c allele to generate a frameshift mutation (tm1d), triggering nonsense mediated decay of the deleted transcript[1,3].

In our previous study, we demonstrated the efficient allele conversion facilitated by electroporation of Cre protein in both zygotes and primary cells. The method was proven to be also compatible with in vitro fertilization procedures without toxic effects on zygotes[11]. To advance the recombinase technology toolbox, we sought to expand it with FLP/*FRT*, Dre/*rox*, and Vika/*vox*, systems and establish a straightforward allele conversion protocol based on AAV delivery without the time and animal consumption, associated with multiple crossings.

Persistent difficulties in the synthesis of FLP (specifically Flpe) and Vika proteins led us to explore alternative recombinase delivery methods. We packaged Flpo and Vika CDS under the control of the *CMV* promoter in recombinant Adeno-Associated Virus vectors (AAV). The delivery in the form of AAV (serotype 1) relies on the virus's natural tropism to zygote's membrane moieties, eliminating the need for electroporation[12,13], and circumvents the issues with protein expression, folding, and overall function. Therefore, this approach does not subject the embryos to physical stress.

In this study, we present a new efficient approach to allele conversion using recombinases delivered via AAVs, representing an attractive alternative to electroporation or mRNA pronuclear microinjection. AAV vectors demonstrated high conversion efficiency in both embryos and adult animals, derived from the MuX reporter mouse. For single-allele conversion, all tested recombinases can reach 100% conversion rate, if optimal titer is used. In addition, we introduce a new approach to multi-level conversion of the EUCOMM/KOMP allele. Our method enables the conversion from tm1a to tm1d in a single animal during IVF-based reanimation/rederivation from sperm, resulting in 80% of fully-converted animals with relatively low AAV vector persistence. These results underscore the spatiotemporal potential of AAV SSR-based conversion and emphasize the reduction of animal consumption in line with 3R principles.

Materials and methods
Construction of mRNA template and AAV SSR transfer plasmids
Vectors were constructed using In-Fusion cloning (Takara Bio) method. Oligonucleotides are shown in Table of Primers (supplementary file) and cloning details are available upon request. In brief, codon-optimized Dre sequence was amplified from pCAG-NLS-HA-Dre (Addgene plasmid #51272) with CDS Dre fw and rv primers, and cloned into pET-Cre-6xhis using *NcoI* and *BamHI* restriction sites. The same approach was used with Flpo and Vika sequence. Vika sequence was amplified form pNPK-NLS-Vika (Addgene plasmid #79969) using CDS Vika fw and rv primers. Flpo sequence was amplified form pCAG-Flpo (Addgene plasmid #60662) using CDS Flpo fw and rv primers. Cloning of the corresponding sequences into pET-Cre-6xhis gave rise to pET-Dre-6xhis, pET-Vika-6xhis pET-Flpo-6xhis, respectively.

The Flpo gene is a mammalian codon-optimized version of the Flpe gene (enhanced flippase), which is derived from the original FLP protein isolated from *Saccharomyces cerevisiae*. Modifications to the FLP coding sequence produced an 'enhanced', more thermostable FLP enzyme, Flpe. Flpe works more efficiently in mice and mammalian cells, but it is generally less efficient than Cre-lox. Further sequence changes to the Flpe recombinase produced 'Flpo', a mouse codon-optimized version whose recombination efficiency approaches that of Cre. In this work, we used molecules/vectors carrying the mammalian codon-optimized variant of the Flpe gene (Flpo), and therefore, we refer to it as such[14]. To clone pAAV-CMV version of recombinase plasmids using In-Fusion cloning (Takara Bio) method, codon-optimized Cre sequence was amplified from pCAG-Cre-IRES2-GFP (Addgene plasmid #26646) with AAV CDS Cre fw and rv primers and cloned into pAAV-GFP Control Vector (AAV-400, Cellbiolabs Inc, San Diego, CA) using *EcoRI* and *XbaI* restriction sites. Similarly, but using *EcoRI* and *HindIII* restriction sites, pAAV CMV-Dre, pAAV CMV-Vika and pAAV CMV-Flpo were constructed. Dre sequence was amplified from pET-Dre-6xHis using AAV CDS Dre fw and rv. Vika sequence was amplified from pET-Vika-6xHis using AAV CDS Vika fw and rv. Flpo sequence was amplified from pET-Flpo-6xHis using AAV CDS Flpo fw and rv. Sequence of used primers is available in Table of Primers (supplementary file). All assembled plasmids generated in this study will be deposited in the Addgene repository to facilitate their availability to the scientific community.

pET-Cre-6xHis was a gift from David Liu (Addgene plasmid #62939). pAAV-EF1a-Flpo was a gift from Karl Deisseroth (Addgene plasmid #55637, Addgene viral prep #55637-AAV1). pCAG-NLS-HA-Dre (Addgene plasmid #51272) was a gift from Pawel Pelczar. pCAG-Flpo (Addgene plasmid #60662) was a gift from Massimo Scanziani. pNPK-NLS-Vika (Addgene plasmid #79969) was a gift from Frank Buchholz. pCAG-Cre-IRES2-GFP (Addgene plasmid #26646) was a gift from Anjen Chenn.

Production of adeno-associated virus and titration
The production of recombinant AAV SSR vectors [AAV-CMV-Cre (Cre), AAV-CMV-Dre (Dre), AAV-CMV-Flpo (CFlpo), AAV-CMV-Vika (Vika)] involved steps as follows: First, 293AAV cells (AAV-100, Cellbiolabs) were cultivated in DMEM (Sigma Aldrich, D6429) supplemented with 10% fetal bovine serum (Invitrogen, 10437028) and 1% penicillin/streptomycin (GIBCO, 15140122) at 37 °C with 5% CO_2. To generate recombinant AAV vectors, the 293AAV cells underwent a triple transfection process with a Helper plasmid (340202, Cellbiolabs), Rep/

2

80

Cap plasmid pAAV2/1 (Addgene plasmid #112862), and a transfer plasmid. The transfection was carried out using polyethylenimine (PEI) (Polysciences, 24765-1) following established protocols[15]. Purity of AAV particles was assessed by Ag-staining using the PierceTM Silver Stain Kit (24612, Thermo Scientific) following the manufacturer's protocol. The DNase-resistant viral genomic titers were determined via qPCR using universal AAV ITR primers (AAV ITR fw and rv) in combination with LightCycler® 480 SYBR Green I Master (Roche) and the LightCycler® 480 System (Roche) following the method described in Aurnhammer[16]. pAAV-CMV-Cre plasmid was used to prepare a standard curve for the absolute quantification of AAV genome copies of purified samples.

Production of mRNA

Cre, Dre, Flpo, and Vika mRNA were transcribed and tailed using the mMessage mMachine Kit (AM1344, Invitrogen) following the manufacturer's protocol from pET-Cre-6xHis (Addgene plasmid #62939), pET-Dre-6xHis, pET-Flpo-6xHis, pET-Vika-6xHis plasmids after *XhoI* (ER0692, Thermo Scientific) linearization. Further purification was performed using the MEGAclearTM Kit (AM1908, Invitrogen) following the manufacturer's protocol. To assess the efficiency of tailing and mRNA integrity, non-tailed and polyA tailed mRNA were separated on a hydrogen peroxide agarose gel following the method described by Pandey and Saluja[17].

Production of Cre and Dre proteins

The Cre (5000 ng/µL) protein was expressed and purified as described by Jenickova[11]. The Dre-6xHis fusion protein was heterologously expressed in *E. coli* BL21-Codon Plus (DE3)-RIPL cells. Briefly, cells were grown in the LB medium supplemented with 0.2% glucose (Sigma) and 100 µM Ampicillin (Sigma) at 37 °C overnight. The following day, cells were diluted 1000-fold into LB medium + Ampicillin and cultured to the OD600 = 0.6 at 37 °C. Subsequently, incubation temperature was lowered to 18 °C, and the recombinant protein expression was induced by the addition of 0.5 mM IPTG; cultivation continued at 18 °C overnight. After the incubation period, cells were harvested by centrifugation (5500G, 12 min) and resuspended in the lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 10% glycerol, 20 mM imidazole, pH 8.0) supplemented with EDTA-free protease inhibitor cocktail (Roche Diagnostics GmbH, Mannheim, Germany). Cells were disrupted using a high-pressure homogenizer (Avestin) and the lysate cleared by centrifugation steps at 7200G (15 min) and 40,000G (30 min). The supernatant was loaded onto a Ni–NTA column (Ni–NTA Superflow, IBA, Germany), washed with the lysis buffer and the Dre-6xHis fusion eluted with the lysis buffer supplemented with 200 mM imidazole. To minimize the presence of contaminating DNA, the pooled elution fractions were mixed with the equal volume of the HA buffer (100 mM sodium phosphate, 300 mM NaCl, 10% glycerol, 100 mM imidazole, pH 8.0) and loaded onto a hydroxyapatite column (BioRad). The flow-through fraction comprising the Dre-6xHis fusion was pooled, dialyzed against PBS and concentrated to 5.2 mg/mL.

Animals

C57Bl/6NCrl (donor) females between the ages of 3–5 weeks were purchased from Charles River. Rosa26-VFRL-EGFP (also known as Multi-reporter mouse line for loci-of-recombination MuX) strain was imported from Max Planck Institute of Molecular Cell Biology and Genetics (Dresden). Crl:CD1 (ICR) surrogate females were purchased from Charles River.

Ethics and compliance statement

Mice were bred in our specific pathogen-free facility (Institute of Molecular Genetics of the Czech Academy of Sciences; IMG). All animals were handled in accordance with the Guide for the Care and Use of Laboratory Animals, approved by the Animal Care and Use Committee of the Academy of Sciences of the Czech Republic. All experiments in this study were carried out in compliance with the ARRIVE guidelines and the laws of the Czech Republic. Animal protocol (93/2020) was approved by the Resort Professional Commission for Approval of Projects of Experiments on Animals of the Czech Academy of Sciences, Czech Republic.

All mice were euthanized using carbon dioxide (CO_2) inhalation followed by cervical dislocation. CO_2 inhalation was performed at a flow rate that displaced 30% of the chamber volume per minute to minimize distress.

Mouse MuX zygote isolation and AAV treatment

The detailed procedure for hormonal stimulation in mice was previously outlined by Luo[18] and is considered a standard practice at transgenic core facilities. In summary, matured C57Bl/6NCrl females aged 28–35 days underwent hormonal stimulation with 5 IU PMSG. After 46 h, they received an additional 5-IU hCG injection and were subsequently mated with homozygote MuX adult males. Plugged females were selected for zygote isolation from the oviducts using M2 medium. The zygotes were briefly cultured in a drop of EmbryoMax KSOM media, covered with paraffin oil, and placed in an incubator at 37 °C with 5% CO_2.

For zygotes treatment, a specific titer of AAV SSR vectors was added to KSOM media and incubated overnight at 37 °C with 5% CO_2. The next day, 2-cell embryos were transferred to fresh KSOM (without virus) media and later into Crl:CD1(ICR, Charles River) surrogate females or further cultivated until the blastocyst stage (E4.5).

Mouse zygote electroporation and pronuclear microinjection

Zygotes were electroporated using the NEPA21 type II machine (NEPAGENE) and an electroporation slide (CUY501P1-1.5). Only zygotes with two visible pronuclei were selected for electroporation. The electroporation process involved the use of M2 and OptiMEM medium. The Cre or Dre proteins were diluted to a final concentration of 1000 ng/µL in OptiMEM medium prior to electroporation.
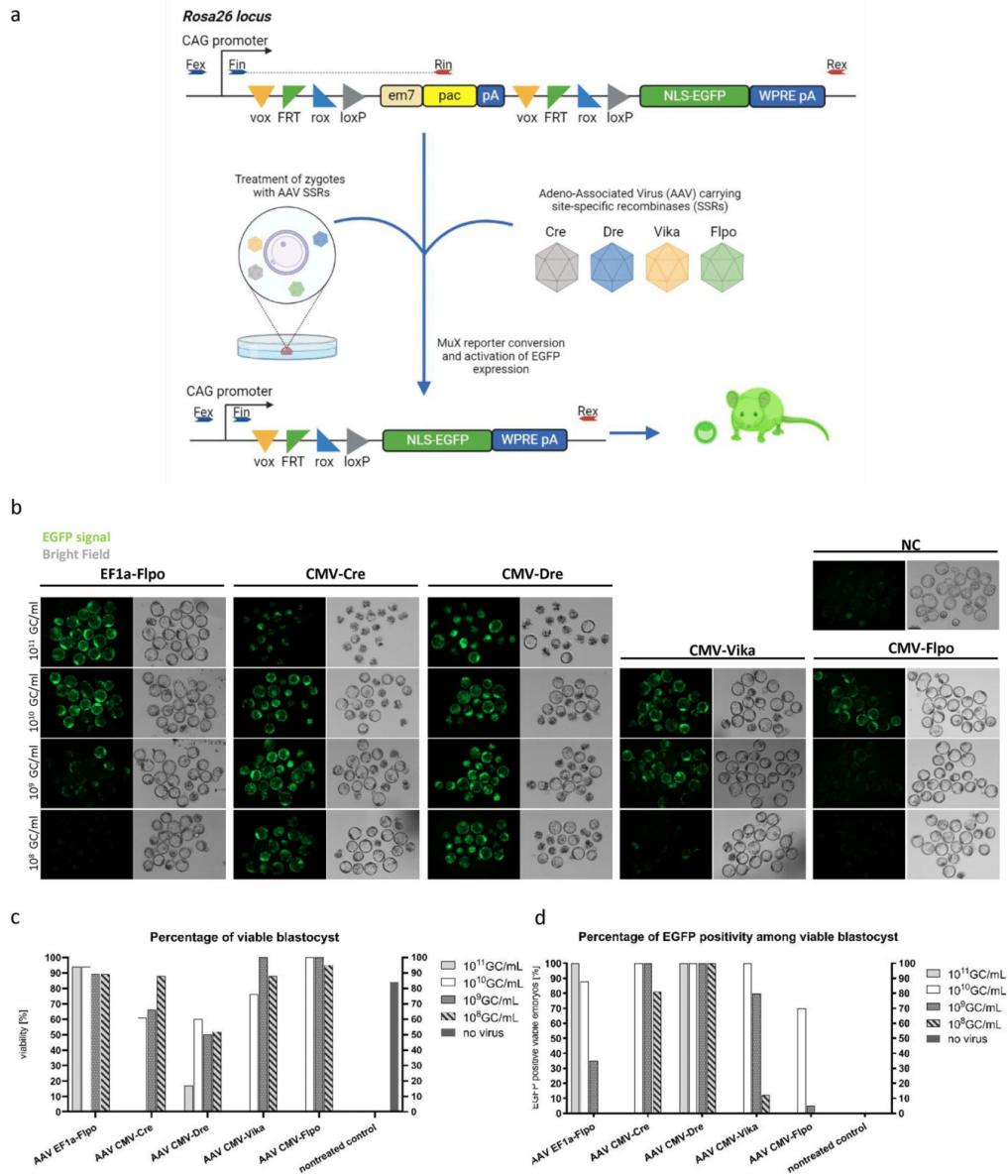
Microinjection of SSR mRNA was performed under the microinjection microscope (Leica DMI6000 B) equipped with FemtoJet 4i microinjector (Eppendorf). Only zygotes with visible pronuclei were selected for microinjection. Each mRNA was injected in the concentration 100 ng/μL.

Mouse in vitro fertilization and AAV treatment with Thawed sperm of tm1a strains

The protocol for mouse IVF has been described previously by Luo[18]. In our study, tm1a heterozygous sperms were utilized in conjunction with C57Bl/6N oocytes. The sperm was frozen previously according to the protocol by Takeo and Nakagata[19].

Briefly, C57Bl/6NCrl females (28–35 days old) were hormonally stimulated with 5-IU PMSG, followed by 5-IU hCG 50 h later. Fourteen to fifteen hours after the hCG application, females were sacrificed, and cumulus-oocyte

◄**Fig. 1.** Titration of AAV SSRs on MuX derived zygotes. (**a**) Conversion of Rosa26-VFRL-EGFP (MuX) reporter using AAV Cre, Dre, Vika and Flpo vector: AAV vectors deliver SSRs into MuX derived zygotes, resulting in conversion of the reporter cassette and activation of EGFP expression. Fully-converted embryos and animals show expression of the reporter in all cell types and tissues. Listed SSRs, Cre, Dre, Vika, Flpo recognize and recombine corresponding specific sites *loxP, rox, vox, FRT*, respectively. Fin and Rin detect the presence of the *pac-pA* cassette preventing *EGFP* expression. Fex and Rex primers amplify wild-type allele *Rosa26* without the MuX reporter and serve as internal control. (**b**) MuX-derived zygotes were subjected to decreasing concentrations of AAV SSRs, spanning from 1E + 11 to 1E + 8 GC/mL. Viability and the presence of enhanced green fluorescent protein (EGFP) signal at the blastocyst stage, indicative of stop cassette recombination, were evaluated for each condition. AAV CMV-Vika and CMV-Flpo treatments at a concentration of 1E + 11 GC/mL are not included due to inadequate vector titers. NC: non-treated control. The scale bar (100 μm) in the image of the non treated control serves as a universal reference applicable to all images within this scheme. Table 1 provides an overview of treatment efficiency based on the observed EGFP signal and embryo viability for each tested condition. (**c**) Graph depicting percentage of blastocyst viability based on used titer. (**d**) Percentage of EGFP blastocysts among viable blastocyst. Graphs (**c**) and (**d**) are based on Tab.1. Blastocysts from each treated group were genotyped, see Supplementary Fig. S1 (supplementary data).

complexes were isolated into a fertilization drop (90 μL of high calcium human tubal fluid (HTF) medium with 1 mM ʟ-glutathione (GSH; Sigma-Aldrich)) containing AAV SSR at a concentration of 5E + 10 GC/mL.

Thawed/extracted sperm was capacitated in MBCD medium for 15–20 min. The capacitated sperm was then transferred into the fertilization drop and incubated with oocytes at 37 °C and 5% CO_2 atmosphere for 4 h. Residual sperm cells were washed out, and fertilized oocytes were transferred into the incubation drop of HTF without GSH (total volume 150 μL) enriched by AAV SSR at a concentration of 5E + 10 GC/mL. Fertilized oocytes were incubated with the virus overnight at 37 °C and 5% CO_2 atmosphere.

The next day, 2-cell stage embryos were collected and transferred into the oviduct of a pseudo-pregnant surrogate female. For tm1a-to-tm1d conversion, viable 2-cell stage embryos were transferred into HTF(-GSH) with the AAV CMV-CRE vector (titer: 1E + 10 GC/mL) and incubated for 4 h or until embryo transfer. The whole procedure is summarized in Fig. 3b.

Genotyping
Blastocysts were collected and lysed in lysis buffer (50 mM Tris HCl pH 8, 1 mM EDTA, 0.5% Triton X-100) with proteinase K (R75282, P-LAB) for 30 min at 55 °C; then inactivated at 95 °C for 15 min.

Tissue samples (tail/ear) of all mentioned strains were lysed in lysis buffer (100 mM Tris HCl pH 8, 200 mM NaCl, 5 mM EDTA, 10% SDS) with proteinase K (R75282, P-LAB) overnight at 55 °C, then inactivated at 95 °C for 15 min. Lysates were further processed by phenol:chlorophorm based isolation using Phenol/Chloroform/Isoamylalcohol reagent (A156.2, ROTI) following the manufacturer's protocol.

Multi-reporter mouse line for loci-of-recombination (X) embryos/mice were genotyped using a two primer pairs reaction. The first pair (Fex and Rex) amplifies the *Rosa26* locus without MuX reporter cassette (wild-type allele) with a corresponding product band at 466 bp, the second pair (Fin and Rin) amplifies the non-converted form of the MuX allele generating a 643 bp product. To detect recombination of the MuX allele with these primers, we always used heterozygote animals in the final analysis in order to distinguish the presence of both alleles (wild-type and MuX allele).

All EUCOMM/KOMP converted strains were genotyped for the presence of the wild-type allele, tm1 allele, tm1c allele, and residual *LacZ* cassette. In the case of *Ube2t* and *C4bp* Cre-converted animals, genotyping focused on the tm1d allele too. Further, all AAV SSR treated mice were analyzed for potential AAV genome persistence (episome/genome integration). The sequence of the used primers is available in the Table of primers (supplementary file).

For detailed description of the genotyping of EUCOMM/KOMP strains and the AAV persistence assay protocol, see the Supplementary Method file and Supplementary Fig. S26. Selected founders for each strain were further bred with wild-type C57Bl/6NCrl animals. Genotyping product of these founders were subjected to Sanger sequencing, data available in supplementary file: Supplementary Data Sanger Sequencing of Founder Mice.

Tissue samples imaging
Organs were captured with a Zeiss Axio Zoom V.16 microscope (200 ms exposure; intestine for 100 ms), using a 450–490 nm excitation wavelength and a 500–550 nm emission wavelength, at a total magnification of 11.2×. E4.5 embryos were captured with an exposure time of 400 ms, using a 480 nm excitation wavelength and a 509 nm emission wavelength, at a total magnification of 179.2×. All images were processed in the ZEN 3.0 (blue edition) software.

Results
Optimal viral titer for efficient AAV-mediated conversion with reduced toxicity in reporter mouse zygotes
Heterozygote MuX reporter mice (Fig. 1a) were used to evaluate the delivery efficiency and potential toxicity of AAV-packaged recombinases on zygotes. We assessed recombination efficiency via imaging of EGFP signal, and toxicity based on the calculation of viable embryos. Treated blastocysts were collected and genotyped for the presence or absence of the *pac-pA* cassette (Supplementary Fig. S1).

83

AAV vector	Titer (GC/mL)	EGFP positive blastocysts		Viable blastocysts		Total no. of blastocysts
EF1a-Flpo	1E + 11	18	100.00%	18	94.70%	19
	1E + 10	16	88.80%	18	94.70%	19
	1E + 09	6	35.30%	17	89.50%	19
	1E + 08	0	0.00%	17	89.50%	19
CMV-Cre	1E + 11	0	0.00%	0	0.00%	18
	1E + 10	11	100.00%	11	61.10%	18
	1E + 09	12	100.00%	12	66.60%	18
	1E + 08	13	81.30%	16	88.80%	18
CMV-Dre	1E + 11	3	100.00%	3	17.60%	17
	1E + 10	9	100.00%	9	60.00%	15
	1E + 09	9	100.00%	9	50.00%	18
	1E + 08	9	100.00%	9	52.90%	17
CMV-Vika	1E + 10	13	100.00%	13	76.50%	17
	1E + 09	16	80.00%	20	100.00%	20
	1E + 08	2	12.51%	16	88.80%	18
CMV-Flpo	1E + 10	14	70.00%	20	100.00%	20
	1E + 09	1	5.00%	20	100.00%	20
	1E + 08	0	0.00%	19	95.00%	20
Negative control	-	0	0.00%	16	84.20%	19

Table 1. Titration of AAV SSRs on MuX derived zygotes.

Distinct viral titers of each AAV recombinase were applied to zygotes, starting at 1E + 11 genomes per milliliter (GC/mL) and descending to 1E + 8 GC/mL. For AAV CMV-Flpo (CFlpo) and Vika, the highest concentration was limited to 1E + 10 GC/mL due to the lower yield of virus titer during production compared to Cre and Dre (as can be referenced in Supplementary Fig. S24).

Despite their substantial conversion efficiency, the high titers of Cre and Dre had a detrimental impact on the development of treated embryos, as depicted in Fig. 1b–d, and quantified in Table 1. The highest conversion efficiency with AAV Vika and CFlpo vectors was observed at a concentration of 1E + 10 GC/mL.

Ready-to-use AAV1 particles carrying the EF1a-Flpo (EFlpo) were used to enhance conversion efficiency with Flpo. This purchased virus allowed us to use 1E + 11 GC/mL titer and assess the efficiency-to-toxicity ratio of the treatment as shown in Table 1. Overall, both AAV Flpo vectors show low toxicity towards treated embryos. Application of commercial EFlpo proved that the efficiency of conversion depends on titer (Table 1, Fig. 1b–d).

Titration screening in blastocysts showed that a viral concentration of 1E + 10 GC/mL yields the optimal efficiency-to-toxicity ratio for CMV-Cre, CMV-Dre, and CMV-Vika, as shown in Table 1 and Fig. 1b–d. In contrast to the Cre or Dre vectors, for Flpo vectors, particularly EFlpo, a higher titer enhances conversion efficiency without a negative impact on the viability of embryos (Fig. 1b–d). Treated blastocysts were collected and genotyped to confirm conversion on the genomic level (Supplementary Fig. S1).

AAV-SSRs have minimal impact on embryo development and show reduced AAV genome persistence potential

After determining the optimal titer for efficient allele conversion with AAV recombinases, we conducted experiments to assess whether AAV treatment at relatively high titers (1E + 10 GC/mL) had any adverse effects on late-stage embryo development following embryo transfer into surrogate females. Utilizing the identified optimal titers for AAV recombinases, we achieved a remarkable 100% efficiency of conversion with AAV EF1a-Flpo, AAV CMV-Cre, -Dre, and -Vika vectors. Notably, in the case of the AAV CMV-Flpo vector, we observed a 92.6% conversion rate among the born animals (total of 27), with 2 (7.4%) exhibiting mosaicism and 2 (7.4%) animals without recombination (Fig. 2c, Table 2). To increase the efficiency of Flpo conversion, we used the EF1a-Flpo vector at 1E + 11 GC/mL titer to convert MuX embryos. Using the EFlpo vector, we reached 100% conversion rate. Interestingly, both CFlpo and EFlpo treated groups produced the highest number of born animals, as shown in Table 2. These findings indicate that AAV-Flpo treatment does not interfere with the viability of embryos in later stages of development, and treated zygotes have the potential to give rise to fully converted animals. We selected a fully-converted male from each AAV treated group and analyzed EGFP signal in liver, lung, skin (ear), heart, intestine and testes. All converted animals were EGFP positive in analyzed organs (Fig. 2a). The remaining animals of the treatment group underwent a biopsy of fingers, which were used to detect the EGFP signal to confirm correspondence to the genotype (Supplementary Figs. S2–S6). Based on the genotype and phenotype analysis, we quantified the number of fully-converted, partially-converted (mosaics) and non-converted animals within each treated group (Table 2, Fig. 2c).

Additionally, we conducted an analysis of treated animals for the persistence of the AAV-based recombination system in either an integrative or episomal state. Specifically, we genotyped converted mice for the AAV genome
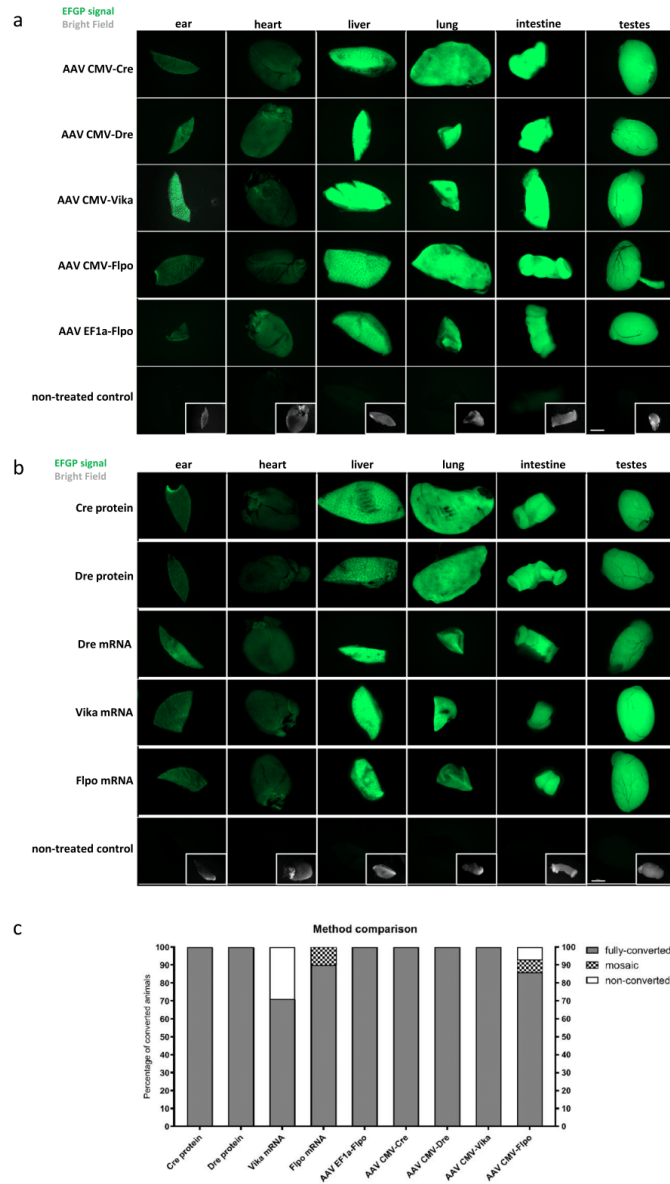
84

Fig. 2. Comparison of conversion efficiency using SSR protein or mRNA with AAV delivered SSRs. MuX heterozygote zygotes underwent conversion by treatment with either AAV or protein/mRNA. (**a**) Imaging of EGFP signal in fully-converted animals by AAV-SSR vectors (Cre, Dre, Vika and CMV-Flpo, EF1a-Flpo). (**b**) Imaging of EGFP signal in fully-converted animals by electroporated Cre or Dre proteins, and microinjection of Vika or Flpo mRNA molecules. (**c**) Graph depicting recombination efficiency throughout delivery methods, based on Table 2 data. The scale bar (2000 μm) in the image of non treated testes serves as a universal reference applicable to all images within this figure.

Delivery	Protein		mRNA		AAV				
Molecule/construct	Cre	Dre	Vika	Flpo	Cre	Dre	Vika	CFlpo	EFlpo
Transferred embryos	91	60	65	60	87	164	82	97	156
Born mice (*)	20 (22%)	13 (21.6%)	7 (10.8%)	10 (16.7%)	16 (18.4%)	13 (7.9%)	8 (9.8%)	27 (27.8%)	36 (23.1%)
Fully recombined	20 (100%)	13 (100%)	4 (57.1%)	9 (90%)	16 (100%)	13 (100%)	8 (100%)	23 (85.2%)	36 (100%)
Mosaics	0	0	1 (14.3%)	1 (10%)	0	0	0	2 (7.4%)	0
Non-recombined	0	0	2 (28.6%)	0	0	0	0	2 (7.4%)	0
AAV genome persistence rate	–	–	–	–	1 (6.3%)	2 (15.4%)	1 (12.5%)	1 (3.7%)	4 (11.1%)

Table 2. Comparison of conversion efficiency using SSR protein or mRNA and AAV delivered SSRs. *Percentage of born pups from transferred embryos. The table presents a summary of genotype and phenotype data for animals in the respective treatment groups, as depicted in supplementary data Supplementary Figs. S2–S10.

of the corresponding SSR. The analysis revealed that AAV SSRs have the potential to persist in treated animals at an average rate of 9.8%. For instance, with CMV-Cre, CMV-Vika and CMV-Flpo vectors, we detected residual AAV genome presence in one animal for each vector out of more than 8 animals (Supplementary Figs. S2c,d, S4c,d and S5c,d). In the case of CMV-Dre we detected 2 animals positive for AAV CMV-Dre genome and in EF1a-Flpo treated animals, 4 were Flpo positive (Supplementary Figs. S3c,d and S6c,d). A detailed summary is provided in Table 2.

AAV-mediated allele conversion is suitable alternative to other conventional (in vitro) methods

To compare AAV-based SSR delivery with more established methods, MuX reporter was converted by using electroporation of Cre and Dre proteins, and microinjection of Vika and Flpo mRNA. Fully-converted males were dissected and selected organs were sampled to confirm their genotype by the presence of EGFP signal (Fig. 2b). The remaining animals of studied groups were analyzed using finger samples (Supplementary Figs. S7–S10). The genotype and phenotype of treated animals are summarized in Table 2 and compared in Fig. 2c.

When employing protein-based conversion, the conversion rate reached 100% in born animals (Table 2, Fig. 2c). However, this method is not available for Flpo and Vika, as Flpo and Vika proteins could not be produced in sufficient quantities despite the use of various (bacterial, baculovirus, mammalian) expression systems (unpublished data). To compare the AAV system with other systems enabling Flpo- and Vika-mediated conversion, we proceeded with the microinjection of Vika and Flpo mRNA. Both Cre and Dre proteins consistently demonstrated 100% conversion efficiency. Flpo mRNA microinjection exhibited a 100% conversion rate with 10% of animals displaying mosaicism, while Vika mRNA application resulted in a 71.4% conversion rate with 14.3% of animals exhibiting mosaic characteristics (Table 2). Thus, the AAV-based delivery SSR system emerges as a fully functional alternative to select established methods of in vitro allele conversion (Fig. 2c).

AAV-Cre and AAV-Flpo vectors enable single or double conversion during mouse line reanimation

To explore the potential of AAV-Flpo recombinase, we attempted to treat presumptive mouse zygote produced by in vitro fertilization from cryopreserved sperm with the AAV-Flpo recombinase. Our protocol was evaluated on selected EUCOMM/KOMP alleles, converting tm1a (KO first allele, reporter-tagged insertion allele) to tm1c (tm1c: conditional allele, post-Flpo) or even tm1d (deletion allele, post-Flpo and Cre with no reporter). The schematic description of the treatment is depicted in Fig. 3a. To convert Aatk[tm1a(KOMP)Wtsi] and Cdh26[tm1a(KOMP)Wtsi] lines to tm1c form, the AAV CMV-Flpo vector (at 1E + 10 GC/mL) was applied. Despite the anticipated reduction in efficiency at the highest attainable titer, the conversion resulted in 3 tm1c-positive out of 4 Aatk tm1-positive pups, where all 3 pups exhibited complete conversion to the tm1c allele (Supplementary Fig. S11). The conversion of Cdh26 zygotes led to the birth of 7 Cdh26 tm1 cassette positive animals, among which 4 were completely converted and 3 were mosaics (Supplementary Fig. S12).

To maximize the yield of fully converted animals, we tested the commercial AAV EF1a-Flpo vector for recombination in the following lines: Rreb1[tm1a(EUCOMM)Wtsi], Atf2[tm1a(EUCOMM)Hmgu] and Ube2t [tm1a(EUCOMM)Hmgu]. The application of the commercial vector yielded high conversion efficiency, with Rreb1 tm1c conversion occurring in all 6 born tm1-positive animals (Fig. 3i–n). Atf2 tm1c conversion was observed in all 15 born tm1-positive pups, where one pup showed a partial conversion (Supplementary Fig. S13). Ube2t tm1c conversion was confirmed in all 7 tm1-positive animals (Supplementary Fig. S14).

To test the limits of AAV-based delivery, the direct in vitro conversion of tm1a alleles to tm1d was performed. This involved the sequential treatment of Ube2t[tm1a(EUCOMM)Hmgu] and C4bp[tm1a(EUCOMM)Hmgu] embryos with AAV EF1a-Flpo and CMV-Cre vectors, respectively. We achieved tm1d conversion in 4 out of 5 tm1-positive pups for Ube2t (Fig. 3b–h), and 4 out of 5 for the C4bp strain (Supplementary Fig. S15). Selected founders of all converted strains were sequenced to confirm specific conversions (Supplementary Data Sanger sequencing of founder mice).

All treated strains and the number of converted animals are summarized in Table 3. One or two of the fully converted and SSR genome negative founders for each strain were subsequently bred with a C57Bl/6NCrl wild-type counterpart. The progeny inherited the fully converted tm1c allele, for Aatk (Supplementary Fig. S16), Cdh26 (Supplementary Fig. S17), Atf2 (Supplementary Fig. S18), Rreb1 (Supplementary Fig. S19), Ube2t
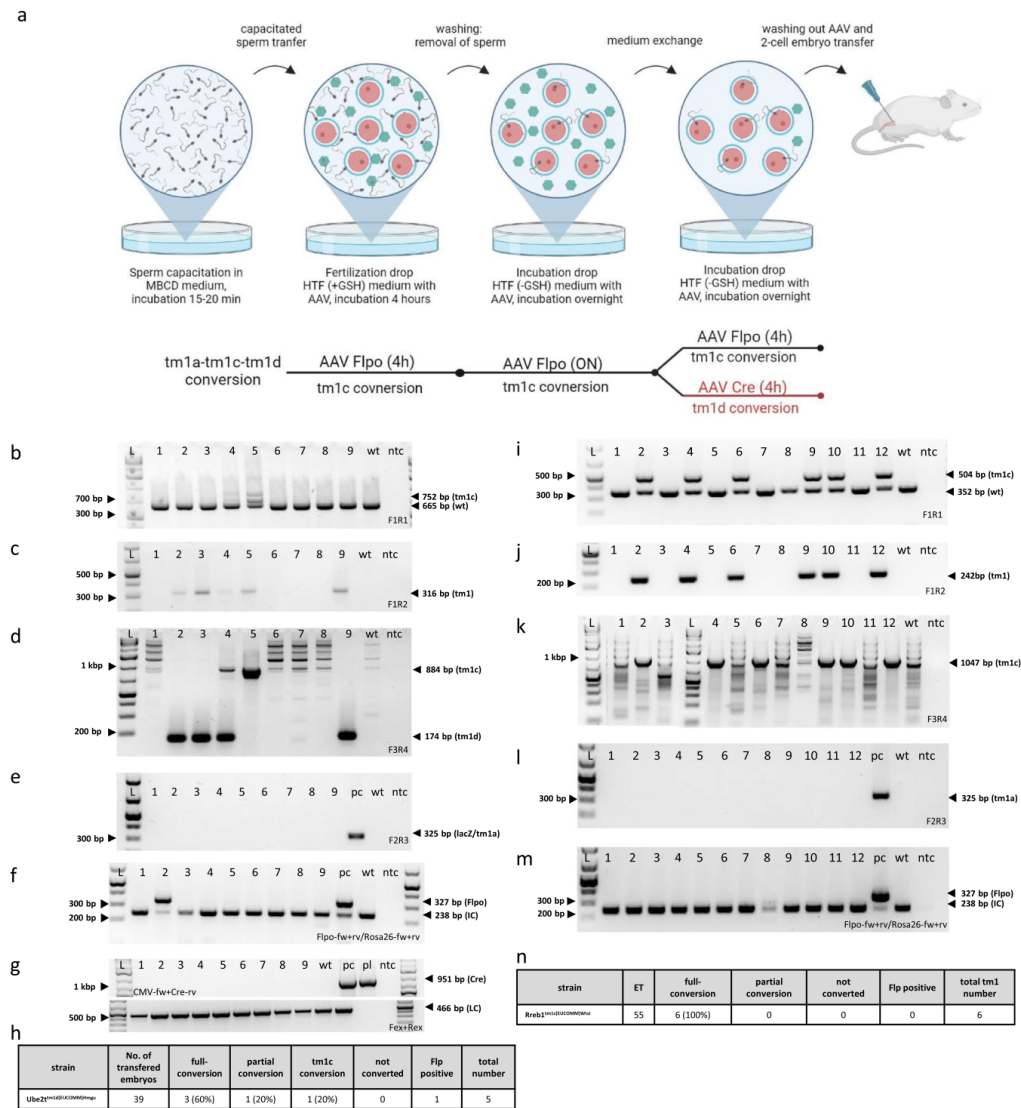
86

Fig. 3. EUCOMM/KOMP allele conversion using AAV Flpo and/or AAV Cre vectors. (**a**) Initially, thawed or extracted sperm undergo capacitation. The capacitated sperm are transferred to a fertilization drop containing AAV SSRs and are incubated. Residual sperm are washed out and fertilized oocytes are transferred to incubation drop with AAV, where they are incubated overnight. On the following day, 2-cell embryos are harvested and transferred into the oviduct of a pseudopregnant recipient female. For tm1a-to-tm1d conversion, viable zygotes, following the overnight treatment, are introduced into an incubation drop with AAV CMV-CRE vector and incubated for 4 h. (**b–e**) Ube2t[tm1d (EUCOMM)Hmgu] genotyping with corresponding primers to detect tm1d allele. (**f,m**) AAV Flpo persistence assay: four primers reaction, Flpo FR primers amplify AAV Flpo genome; internal control (Rosa26 FR primers). wt: C57Bl/6NCrl genomic DNA, pc: positive control—Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (**g**) AAV Cre persistence assay: detection of residual AAV Cre genome, CMV-F and Cre-R primers detect presence of AAV CMV-Cre genome; Fex and Rex primers amplify a DNA loading control reaction; wt:wild-type sample, pc: wild-type DNA and 1 ng/μ of AAV CMV-Cre plasmid, pl: AAV CMV-Cre plasmid, ntc: non-template control. (**h**) Ube2t[tm1d (EUCOMM)Hmgu] genotyping summary. (**i–l**) Rreb1[tm1c (EUCOMMWtsi)] genotyping with corresponding primers to detect tm1c allele. (**n**) Rreb1[tm1c (EUCOMMWtsi)] genotyping summary. L(ladder): GeneRuler 1 kb Plus DNA. Presented figures have been cropped for clarity, complete gel images can be found in Supplementary data Fig. S25.

87

Converted to	No. of tranfered embryos	AAV vector	Full-conversion (tm1c+ and LacZ -)	Partial conversion (tm1c+ and tm1a: LacZ+)	Not converted (tm1a: LacZ+)	Flp positive	Total number (tm1 cassette positive)			GLT*
Aatk tm1c(KOMP)Wtsi	57	CMV-Flpo	3 (75%)	0	1 (25%)	0	4			100%
Cdh26 tm1c(KOMP)Wtsi	88	CMV-Flpo	4 (57.1%)	3 (42.8%)	0	0	7			100%
Rreb1 tm1c(EUCOMM)Wtsi	55	EF1a-Flpo	6 (100%)	0	0	0	6			100%
Atf2 tm1c(EUCOMM)Hmgu	51	EF1a-Flpo	14 (93.4%)	1 (6.6%)	0	0	15			100%
Ube2t tm1c(EUCOMM)Hmgu	56	EF1a-Flpo	7 (100%)	0	0	0	7			100%
Converted to	**No. of tranfered embryos**	**AAV vector**	**Full-conversion (tm1d+ and tm1c-)**	**Partial conversion (tm1d+ and tm1c+)**	**tm1c conversion (LacZ-)**	**Not converted (LacZ+)**	**Flp positive**	**Cre positive**	**Total number (tm1 cassette positive)**	**GLT**
Ube2t tm1d(EUCOMM) Hmgu	39	EF1a-Flpo/ CMV-Cre	3 (60%)	1 (20%)	1 (20%)	0	1	0	5	100%
C4bp tm1a(EUCOMM)Hmgu	54	EF1a-Flpo/ CMV-Cre	4 (80%)	0	1 (20%)	0	1	0	5	100%

Table 3. EUCOMM/KOMP allele conversion using AAV Flpo and/or AAV Cre vectors. germ-line transmission of target allele form in animals positive for tm1 allele. The table provides a summary of genotyping data for the corresponding strain and treatment, as illustrated in Supplementary data Figs. S11–S19.

(Supplementary Fig. S20) or tm1d for Ube2t (Supplementary Fig. S21) and C4bp (Supplementary Fig. S23) without any residual presence of AAV genome. This pioneering strategy underscored the potency of AAV recombinases in the direct ex vivo conversion of tm1a alleles, resulting in rapid conversion to tm1c and/or tm1d in the first generation of transgenic mouse line immediately after fertilization. This approach not only significantly reduces the consumption of animals but also seamlessly aligns with the principles of the 3Rs concept (Replacement, Reduction, Refinement) accepted in the scientific community. All treated strains and numbers of converted animals are summarized in Table 3.

Discussion

Application of complex transgenic alleles, exemplified by EUCOMM/KOMP alleles, often requires multiple recombination events to understand specific biological processes fully. To achieve this level of genetic engineering sophistication, an efficient combinatorial method based on site-specific recombinases (SSRs) is needed. In this context, we introduce a new method utilizing AAV vectors to facilitate such an approach.

As shown with Cre protein[11], we attempted to produce and purify the recombinant Dre, Flpe, and Vika proteins in multiple cellular heterologous systems, including BL21-DE3-RIPL bacterial (BL21-DE3-RIPL), Hi5 insect cells, and HEK293T mammalian cells (HEK293). Nevertheless, Flpe and Vika proteins were not sufficient for experiments in large-scale workflows (unpublished data). To overcome this limitation, SSR coding sequences were packaged in AAV serotype 1 vectors, and their conversion efficiency was compared to established SSR delivery techniques in the form of protein electroporation and/or mRNA microinjection.

Titers of the vectors varied. Cre and Dre containing vectors could be prepared in high-titers, while Vika and CMV-Flpo vectors achieved relatively low-titer (Supplementary Fig. S24). The reason for this discrepancy is unclear. The constructs fall below the AAV packaging limit of 4.7 kb (Cre: 3167 bp, Dre: 3156 bp, Vika: 3159 bp, CMV-Flpo: 3369 bp), thus, the size of the vector should not influence packaging. Despite employing identical procedure for the production and purification protocols, our data suggest that the expression of Vika and Flpo may influence the AAV replication process in a mammalian system. However, it is important to highlight that our data did not indicate a significant presence of empty capsids in the Vika and CMV-Flpo vectors. If such capsids were prevalent, a disproportionately higher amount of protein relative to the GC (genome copy) count would be expected (Supplementary Fig. S24). Therefore, it is more plausible that the efficiency of AAV-SSRs production is negatively affected by the overexpression of Vika and Flpo in mammalian producer cells.

Different AAV SSR titers were applied to MuX heterozygote (het) zygotes to determine the most effective and least toxic viral concentration (Fig. 1b,c). Once the optimal titer was identified, the efficiency and scale of conversion were evaluated in adult mice. Performance of AAV vectors was compared with standard delivery methods, such as electroporation of protein (Cre and Dre), or microinjection of mRNA (Vika and Flpo).

Despite the relatively low-yield in production of CMV-Flpo and Vika vectors (reaching a maximum titer of 7E + 10 GC/mL and 3.3E + 11 GC/mL, respectively), the titration screen identified the minimum effective titers for CMV-Flpo and Vika as 1E + 10 GC/mL and 1E + 9 GC/mL, respectively (Fig. 1b–d). Reduced embryo viability at 1E + 11 GC/mL was observed in the case of Cre and Dre vectors (Fig. 1b,c). The optimal titers were determined based on the viability and conversion efficiency ratio, as quantified in Table 1. However, it is important to note that Cre and Dre vectors can be used at concentrations one order of magnitude lower while still ensuring efficient conversion and low toxicity. In contrast, CMV-Flpo and Vika titration experiments indicated that concentrations below 1E + 10 GC/mL significantly reduce conversion efficiency. A similar trend was observed with the commercial EF1a-Flpo vector. However, unlike Cre and Dre, the highest titer of this vector did not adversely affect embryo viability (Fig. 1b,c).

88

While AAV vectors are generally considered non-integrative vectors, recent evidence has introduced some uncertainty regarding their potential integration following double-stranded breaks (DSB)[20,21]. SSRs, unlike site-specific endonucleases, do not induce open DSB. SSR-mediated recombination represents a more isolated and tightly regulated process compared to the machinery involved in DSB repair[22]. Furthermore, the AAV vectors used in this study do not carry any recombination sites recognized by the given site-specific recombinases. This factor might considerably reduce the likelihood of AAV genome integration during SSR-mediated recombination. Therefore, we hypothesize that the observed persistence of the AAV genome is associated with the episomal state of the AAV genome or spontaneous integration into the genome, rather than with SSR activity[20].

The presence of expression-competent fragments of the AAV vector within the host cells was analyzed, since AAVs are associated with low integrative potential[21]. Our data suggest that all used recombinases in AAV form can persist in the host tissue. Based on AAV persistence assay in MuX mice, the frequency of residual AAV genome is on average 9.8%, which is relatively high by standards for gene therapy (Supplementary Figs. S2–S6). Due to this persistence potential, newborn animals should be screened for Flpo and Cre AAV genomes after AAV-mediated conversions. Based on this selection, there is no need for further back-cross breeding to eliminate the active *Flp* gene from the genome, as is often case with conventional Flpo driver lines.

Notably, the AAV vector can integrate into the genome, frequently at the AAVS1 site, or it can remain in the nucleus in an episomal state[20]. Flpo-positive animals converted by AAV can generate progeny both with and without Flpo integration, as demonstrated in Supplementary Fig. S22. This observation suggests that while AAV-converted animals can test positive for Flpo, the recombinase gene is not necessarily passed to the next generation. This is likely due to the potential episomal state of the AAV vector, mosaic integration in germ cells, or allelic crossover during meiosis. One of the practical applications of the AAV-based SSRs delivery is in vitro fertilization. A single treatment with AAV Flpo or Cre facilitates the direct conversion of the previously mentioned EUCOMM/KOMP alleles without a significant risk of recombinase vector persistence. This approach offers a simple and effective method for Cre or Flpo-mediated conversion. The AAV system not only enables the conversion from tm1a to tm1c, but also provides the opportunity to perform tm1a to tm1d conversion in the whole animal, as shown in Fig. 3b–n. Furthermore, we verified that AAV-mediated conversion is heritably transmitted to the next generation. Since AAV vectors can remain present in either episomal or integrated states, we suggest the preference for AAV SSR-free animals for further work (Supplementary Figs. S16–S23).

We have effectively demonstrated the practical application of AAV SSRs in the context of EUCOMM/KOMP allele conversion, specifically transitioning from the tm1a state to tm1c, and further to tm1d in a single animal (Fig. 3b–h and Supplementary Fig. S15). A distinctive feature of the AAV-based approach is the ability to convert tm1a alleles to tm1d through sequential treatment with two distinct vectors. The current practice for achieving tm1d allele involves converting the tm1a allele to tm1c using either a Flpo driver or mRNA microinjection[23,24]. Subsequent progeny of tm1c mice are then converted either through a Cre driver or mRNA/protein microinjection/electroporation[5,11,23–25]. As a result, the conventional approach requires more time and animals until the final model is established, which can be at odds with the principles outlined in the 3Rs guidelines.

The main objective of this study is to underscore the versatility of AAV SSRs and improve conversion methods employing these enzymes. We focused on the FLP/*FRT* system, whose protein synthesis is complicated, and mRNA microinjection is invasive and less efficient. The utilization of AAV Flpo construct offers an efficient and straightforward tool for effective tm1c conversion without inducing significant adverse effects. Additionally, it enables the potential for combinatorial treatment with AAV Cre vectors to achieve the tm1d allele, resulting in rapid conversion compared to conventional methods. Furthermore, combinatorial treatment enables a reduction in the number of required animals.

Moreover, the use of recombinant AAVs (rAAVs) can be conducted within a biosafety level 1 (BSL-1) facility, provided that the AAV vector does not carry an oncogenic or toxic payload and is produced in a helper virus-free manner. rAAVs loaded with SSRs adhere to the BSL-1 criteria, making the vector readily applicable in laboratories lacking BSL-2 facilities[26]. Furthermore, rAAV vectors can be purchased in a high-purity grade and ready-to-use form, making AAV-mediated conversion a broadly available method. To facilitate the availability of the method, all plasmids generated in this study will be deposited in a non-profit plasmid repository.

Conclusion
In summary, this study highlights the potential of AAV Site-Specific Recombinases (SSRs) for precise allele conversion. Although the dominant Cre/*loxP* system provides a robust framework for achieving spatiotemporal control over alleles through recombination events, Flpo, Dre, and Vika, are not as commonly used for direct ex vivo allele conversion. The delivery of Cre, Dre, Flpo, and Vika recombinases via AAV1 particles resulted in efficient conversion at various target sites while minimizing risk of AAV genome persistence and ensuring the safety of treated embryos. This approach seemlessly aligns with in vitro fertilization (IVF) procedures applied on EUCOMM/KOMP lines. The AAV vectors demonstrate excellent conversion capabilities with a reduced risk of AAV genome persistence. Incorporating these findings, we provide a safe and efficient pathway for allele conversion, with the specific focus on the remarkable potential of AAV Flpo vectors. Our research underscores the versatility and safety of recombinant AAVs within research facilities, irrespective of biosafety level, and highlights AAV Flpo as a standout candidate for efficient and reliable allele conversion.

Data availability
All data generated or analyzed during this study are included in this published article (and its Supplementary Information files). Knock-out first (EUCOMM/KOMP) mouse lines, used in this study, are available in the European Mouse Mutant Archive (EMMA).

89

References

1. Coleman, J. L. J. *et al.* Rapid knockout and reporter mouse line generation and breeding colony establishment using EUCOMM conditional-ready embryonic stem cells: A case study. *Front. Endocrinol. (Lausanne)* **6**, 105 (2015).
2. Loulier, K. *et al.* Multiplex cell and lineage tracking with combinatorial labels. *Neuron* **81**, 505–520 (2014).
3. Skarnes, W. C. *et al.* A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337–344 (2011).
4. Sket, T., Falcomatà, C. & Saur, D. Dual recombinase-based mouse models help decipher cancer biology and targets for therapy. *Cancer Res.* **83**, 2279–2282 (2023).
5. Kim, H., Kim, M., Im, S. K. & Fang, S. Mouse Cre-LoxP system: General principles to determine tissue-specific roles of target genes. *Lab. Anim. Res.* **34**, 147–159 (2018).
6. Karimova, M. *et al.* A single reporter mouse line for Vika, Flp, Dre, and Cre-recombination. *Sci. Rep.* **8**, 1–12 (2018).
7. Karimova, M. *et al.* Vika/vox, a novel efficient and specific Cre/loxP-like site-specific recombination system. *Nucleic Acids Res.* **41**, e37 (2013).
8. Kouvaros, S. et al. A CRE/DRE dual recombinase transgenic mouse reveals synaptic zinc-mediated thalamocortical neuromodulation. *Sci Adv* **9**, (2023).
9. Wang, H. et al. Dual Cre and Dre recombinases mediate synchronized lineage tracing and cell subset ablation in vivo. *J. Biol. Chem.* **298** (2022).
10. Pettitt, S. J. *et al.* Agouti C57BL/6N embryonic stem cells for mouse genetic resources. *Nat. Methods* **6**, 493–495 (2009).
11. Jenickova, I. *et al.* Efficient allele conversion in mouse zygotes and primary cells based on electroporation of Cre protein. *Methods* **191**, 87–94 (2021).
12. Chen, S. *et al.* CRISPR-READI: Efficient generation of knockin mice by CRISPR RNP electroporation and AAV donor infection. *Cell Rep.* **27**, 3780-3789.e4 (2019).
13. Mizuno, N. *et al.* Intra-embryo gene cassette knockin by CRISPR/Cas9-mediated genome editing with adeno-associated viral vector. *iScience* **9**, 286–297 (2018).
14. Kranz, A. *et al.* An improved Flp deleter mouse in C57Bl/6 based on Flpo recombinase. *Genesis* **48**, 512–520 (2010).
15. Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nat. Protoc.* **1**, 1412–1428 (2006).
16. Aurnhammer, C. *et al.* Universal real-time PCR for the detection and quantification of adeno-associated virus serotype 2-derived inverted terminal repeat sequences. *Hum. Gene Ther. Methods* **23**, 18–28 (2012).
17. Pandey, R. & Saluja, D. Hydrogen peroxide agarose gels for electrophoretic analysis of RNA. *Anal. Biochem.* **534**, 24–27 (2017).
18. Luo, C. *et al.* Superovulation strategies for 6 commonly used mouse strains. *J. Am. Assoc. Lab. Anim. Sci.* **50**, 471 (2011).
19. Takeo, T. & Nakagata, N. Reduced glutathione enhances fertility of frozen/thawed C57BL/6 mouse sperm after exposure to methyl-beta-cyclodextrin. *Biol. Reprod.* **85**, 1066–1072 (2011).
20. Deyle, D. R. & Russell, D. W. *Adeno-Associated Virus Vector Integration.*
21. Wolter, J. M. *et al.* Cas9 gene therapy for Angelman syndrome traps Ube3a-ATS long non-coding RNA. *Nature* **587**, 281–284 (2020).
22. Grainge, I. & Jayaram, M. The integrase family of recombinases: Organization and function of the active site. *Mol. Microbiol.* **33**, 449–456 (1999).
23. Ludwig, D. L., Stringer, J. R., Wight, D. C., Doetschman, T. C. & Duffy, J. J. FLP-mediated site-specific recombination in microinjected murine zygotes. *Transgenic Res.* **5**, 385–395 (1996).
24. Wu, Y., Wang, C., Sun, H., LeRoith, D. & Yakar, S. High-efficient FLPo deleter mice in C57BL/6J background. *PLoS One* **4**, 8054 (2009).
25. Luckow, B. *et al.* Microinjection of Cre recombinase protein into zygotes enables specific deletion of two eukaryotic selection cassettes and enhances the expression of a DsRed2 reporter gene in Ccr2/Ccr5 double-deficient mice. *Genesis* **47**, 545–558 (2009).
26. Collins, D. E., Reuter, J. D., Rush, H. G. & Villano, J. S. Viral vector biosafety in laboratory animal research. *Comp. Med.* **67**, 215 (2017).

Acknowledgements

Author contributions

P.N. and I.J. conceived the experiments, cloning work conducted by J.E., P.K. and P.N., production of AAVs conducted by P.N. and I.J., animal work planned and performed by P.K., P.N. I.J. and J.K., production of recombinant proteins C.B. and P.K., P.N. analyzed the results. Writing manuscript, review, editing; P.N., P.K., R.S. Funding acquisition, supervision, administration; RS. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-70853-1.

Correspondence and requests for materials should be addressed to P.N. or R.S.

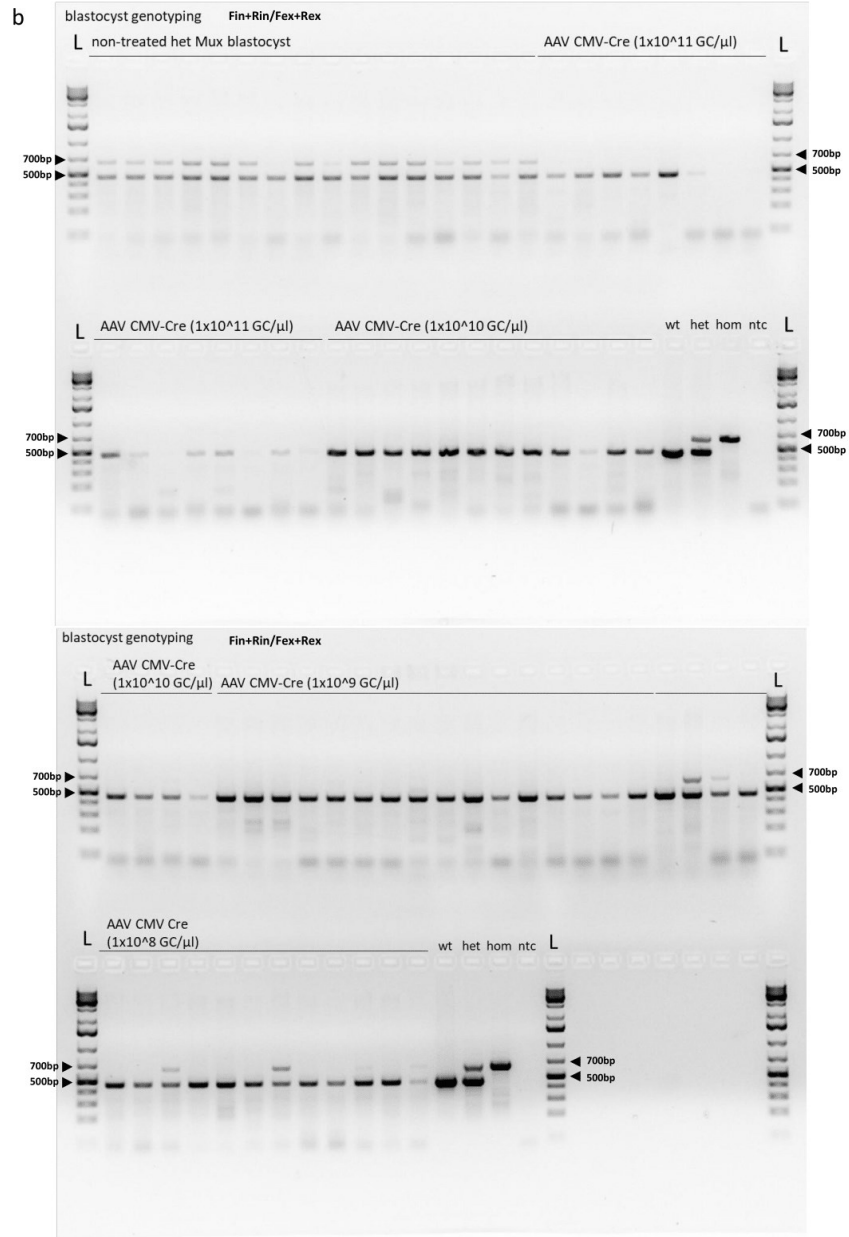**Reprints and permissions information** is available at www.nature.com/reprints.

90

Supplementary figures:

91

a

blastocyst genotyping   **Fin+Rin/Fex+Rex**

blastocyst genotyping   **Fin+Rin/Fex+Rex**

b

blastocyst genotyping   **Fin+Rin/Fex+Rex**

L non-treated het Mux blastocyst   AAV CMV-Cre (1x10^11 GC/µl)   L

700bp
500bp

L AAV CMV-Cre (1x10^11 GC/µl)   AAV CMV-Cre (1x10^10 GC/µl)   wt het hom ntc L

700bp
500bp

700bp
500bp

blastocyst genotyping   **Fin+Rin/Fex+Rex**

   AAV CMV-Cre
L (1x10^10 GC/µl) AAV CMV-Cre (1x10^9 GC/µl)   L

700bp
500bp

700bp
500bp

   AAV CMV Cre
L (1x10^8 GC/µl)   wt het hom ntc L

700bp
500bp

700bp
500bp

c

blastocyst genotyping   **Fin+Rin/Fex+Rex**

AAV CMV-Dre (1x10^11 GC/µl)   AAV CMV-Dre (1x10^10 GC/µl)

AAV CMV-Dre (1x10^10 GC/µl)   AAV CMV-Dre (1x10^9 GC/µl)   wt het hom ntc

blastocyst genotyping   **Fin+Rin/Fex+Rex**

AAV CMV-Dre
(1x10^9GC/µl)   AAV CMV-Dre (1x10^8 GC/µl)   wt het hom ntc

d blastocyst genotyping **Fin+Rin/Fex+Rex**

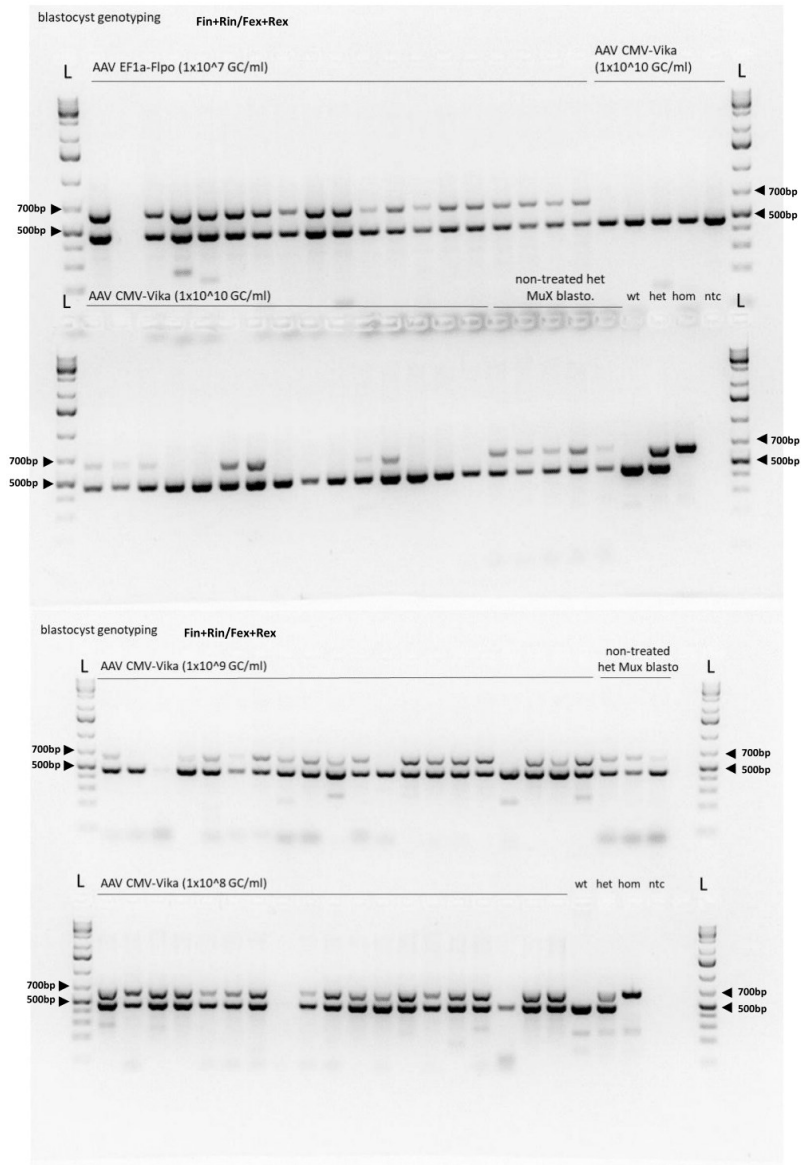**Figure S1: Genotyping of treated heterozygous MuX embryos (E4.5) using Fex/Rex and Fin/Rin primers**

(a) Genotyping of AAV EF1a-Flpo treated blastocysts with corresponding viral titre. (b) Genotyping of AAV CMV-Cre treated blastocysts with corresponding viral titre. (c) Genotyping of AAV CMV-Dre treated blastocysts with corresponding viral titre. (d) Genotyping of AAV CMV-Vika treated blastocysts with corresponding viral titre or non-treated controls. (e) Genotyping of AAV CMV-Flpo treated blastocysts with corresponding viral titre or non-treated controls. Ladder (L): GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue sample without MuX reporter in the genome; het: heterozygote sample of a mouse carrying a MuX allele, hom: homozygote sample of a mouse carrying two MuX alleles; non-treated control: genotyping of 16 random blastocysts from corresponding group in Fig.1. Rosa26 wild type allele without MuX cassette corresponds to a band of 466bp (Fex and Rex primers), non-recombined MuX cassette is indicated by 643bp band (Fin and Rin primers).

Figure S2: Tissue genotyping and imaging of heterozygous MuX mice after AAV CMV-Cre treatment
(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder (L): GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without (deleted) MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, MuX NC: non-recombined MuX heterozygote finger, magnification 11,2x. (c) AAV genome persistence assay: detection of residual AAV Cre genome, CMV-F and Cre-R primers detect presence AAV CMV-Cre genome, band corresponding to 951bp; Fex and Rex primers produce 466bp band as a DNA loading control reaction; wt:wild-type sample, pc: wild-type DNA and 1ng/µ of AAV CMV-Cre plasmid, pl: AAV CMV-Cre plasmid, ntc: non-template control. (d) Summary of genotype and phenotype analysis.
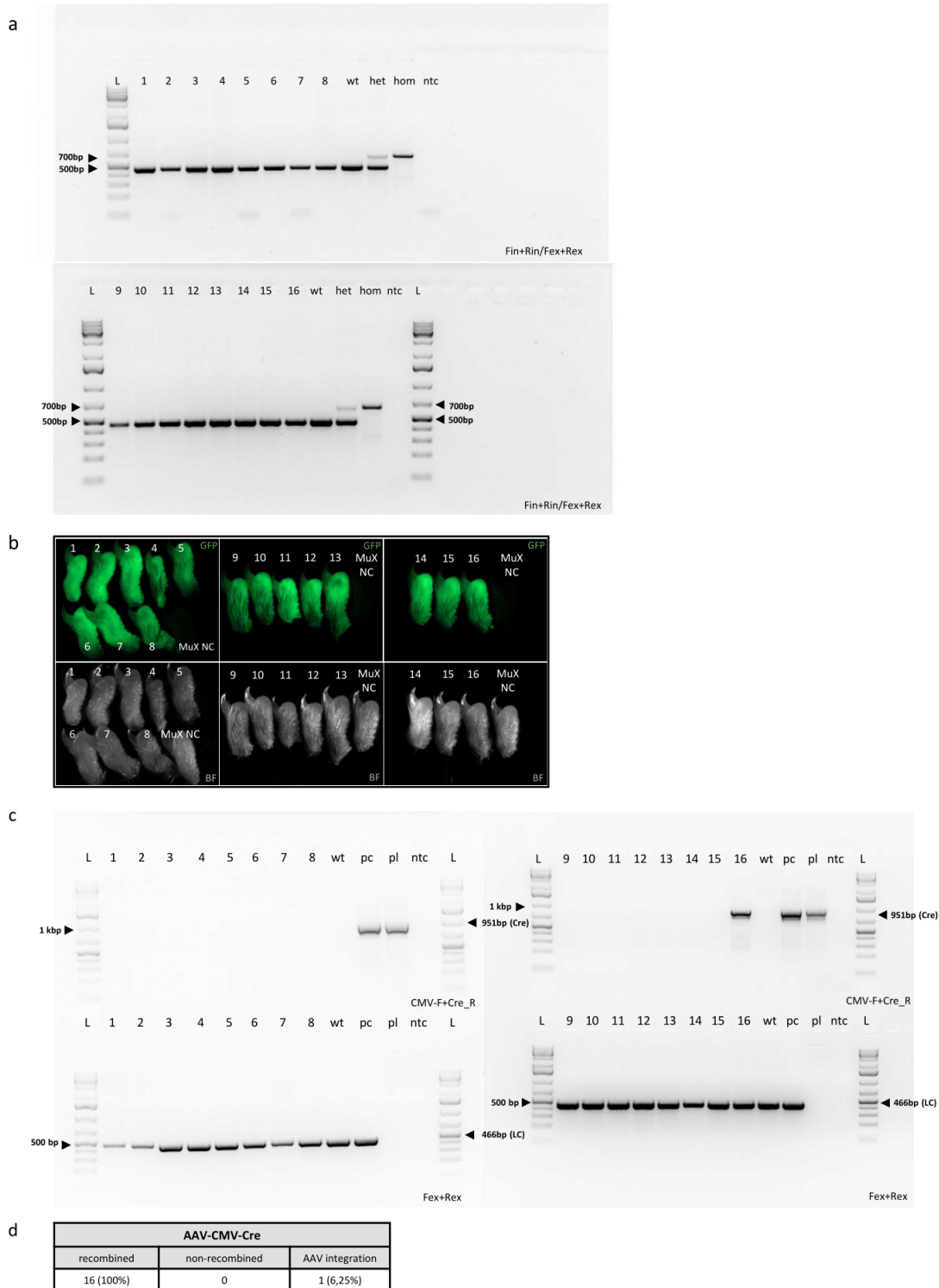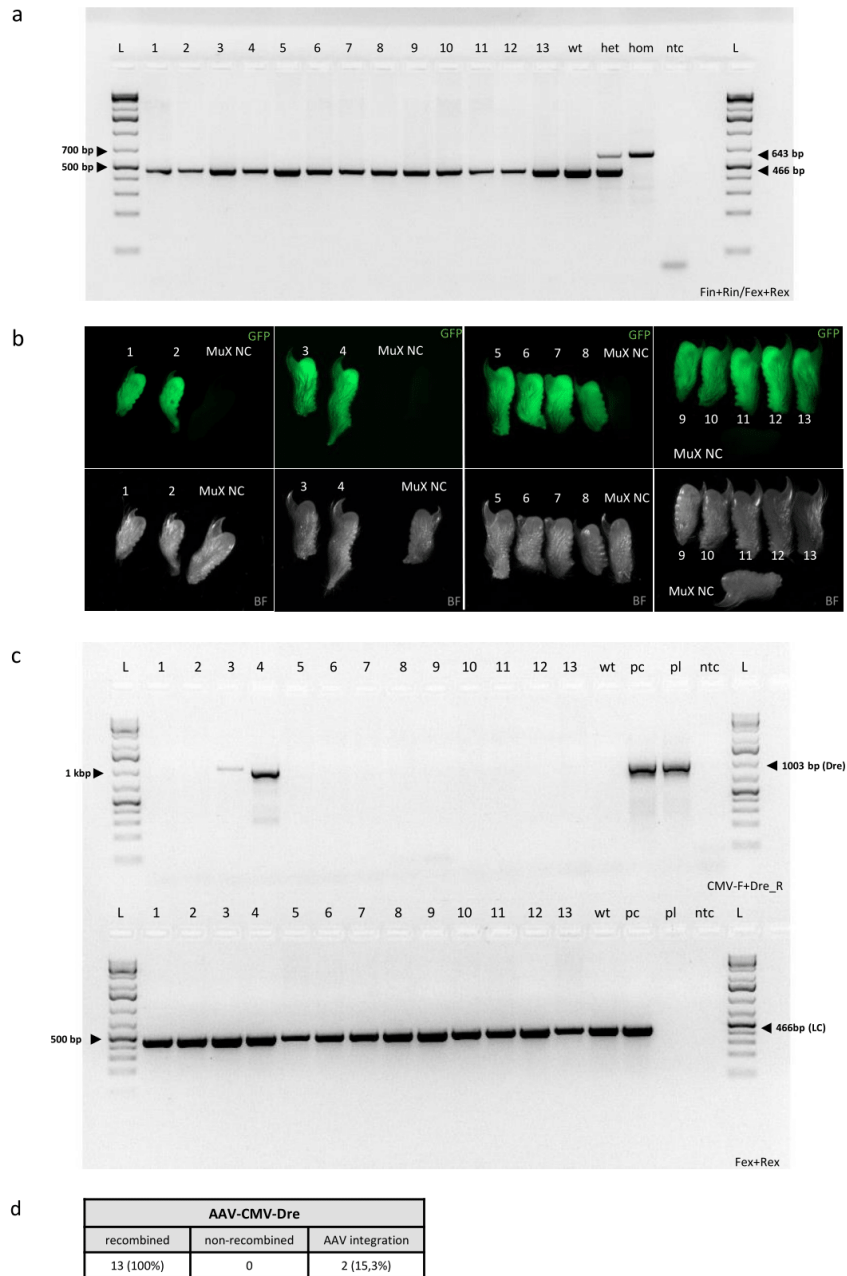
Figure S3: Tissue genotyping and imaging of heterozygous MuX mice after AAV CMV-Dre treatment

(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder (L): GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, MuX NC: non-recombined MuX heterozygote finger, magnification 11,2x (c) AAV genome persistence assay: detection of residual AAV Dre genome, CMV-F and Dre-R primers detect presence AAV CMV-Dre genome, band corresponding to 1003bp; Fex and Rex primers produce 466bp band as a DNA loading control reaction; wt:wild-type sample, pc: wild-type DNA and 1ng/µ of AAV CMV-Dre plasmid, pl: AAV CMV-Dre plasmid, ntc: non-template control. (d) Summary of genotype and phenotype analysis.

a

L 1 2 3 4 5 6 7 8 wt het hom ntc L

700 bp ▶
500 bp ▶

Fin+Rin/Fex+Rex

b

GFP   GFP
1 2 3 MuX NC   4 5 6 7 8 MuX NC

1 2 3 MuX   4 5 6 7 8 MuX
        NC   NC
BF   BF

c

L 1 2 3 4 5 6 7 8 wt pc pl ntc

1 kbp ▶   ◀ 1083 bp (Vika)
500 bp ▶

L 1 2 3 4 5 6 7 8 wt pc pl ntc   CMV-F+Vika_R

500 bp ▶   ◀ 466bp (LC)

Fex+Rex

d

AAV-CMV-Vika		
recombined	non-recombined	AAV integration
8 (100%)	0	1 (12,5%)

Figure S4: Tissue genotyping and imaging of heterozygous MuX mice after AAV CMV-Vika treatment

(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder (L): GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, MuX NC: non-recombined MuX heterozygote finger, magnification 11,2x. (c) AAV genome persistence assay: detection of residual AAV Vika genome, CMV-F and Vika-R primers detect presence AAV CMV-Vika genome, band corresponding to 1083bp; Fex and Rex primers produce 466bp band as a DNA loading control reaction; wt:wild-type sample, pc: wild-type DNA and 1ng/µ of AAV CMV-Vika plasmid, pl: AAV CMV-Vika plasmid, ntc: non-template control. (d) Summary of genotype and phenotype analysis.

d

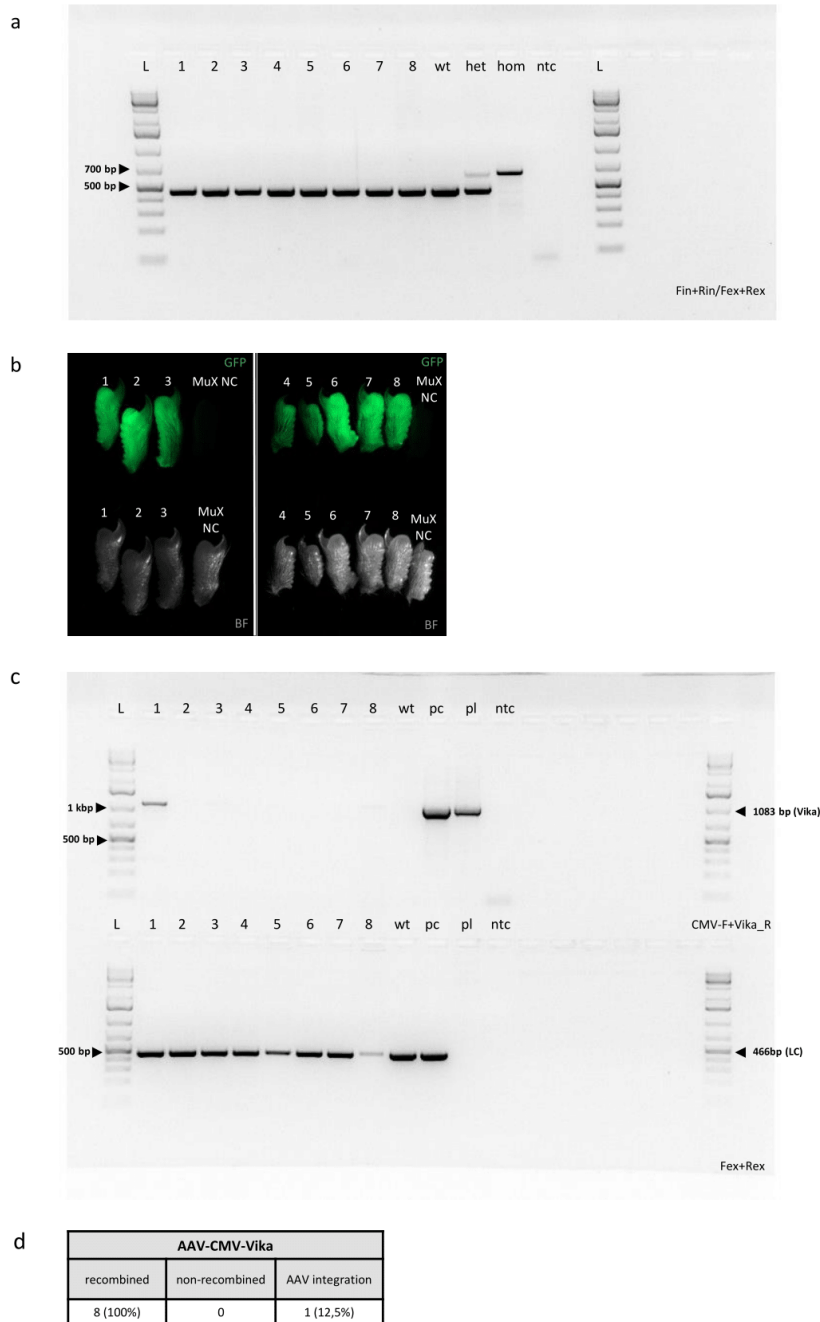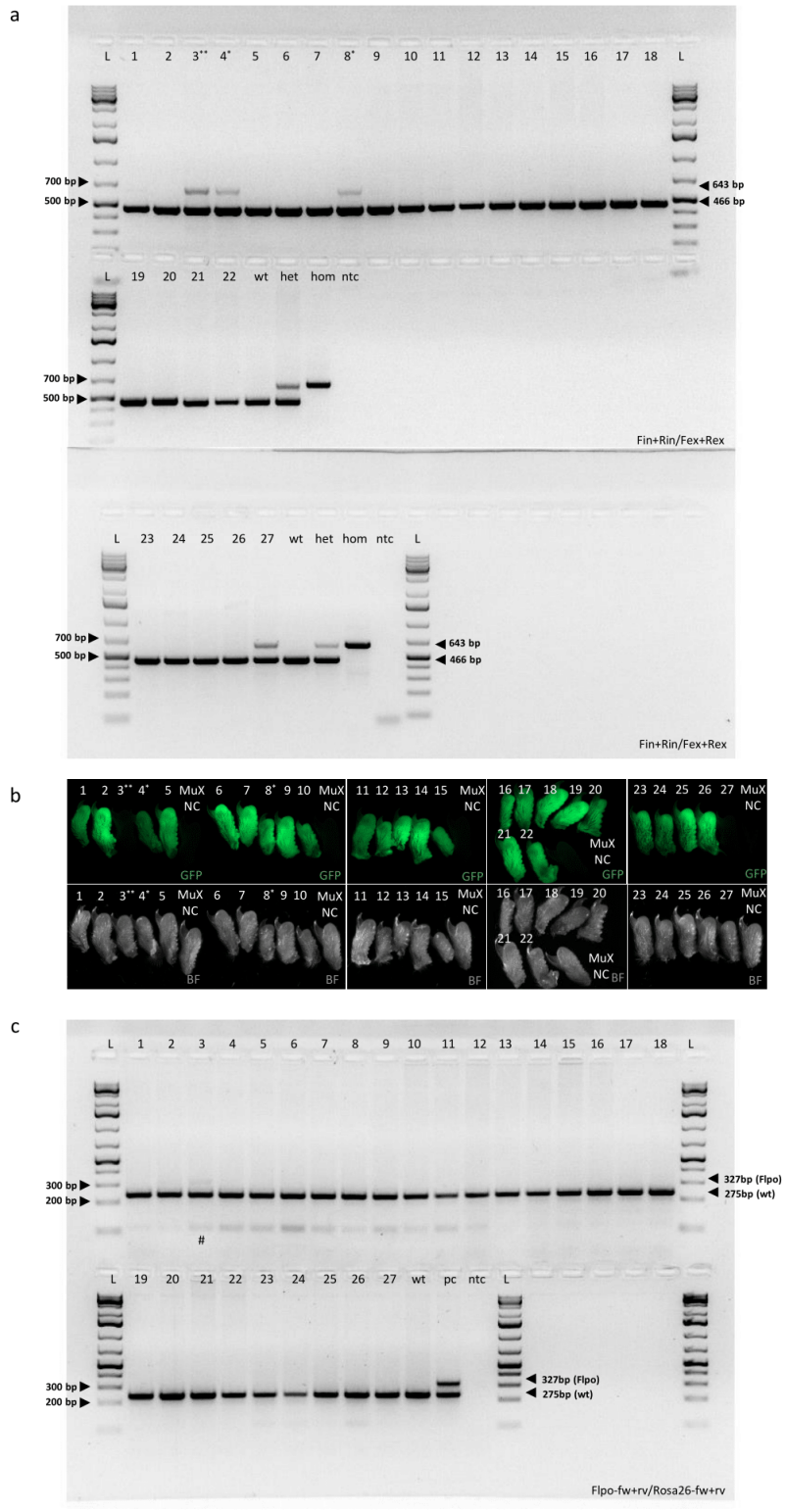| AAV CMV-Flpo			
recombined	*mosaic	**non-recombined	#AAV integration
23 (85,2%)	2	2	1 (3,7%)

Figure S5: Tissue genotyping and imaging of heterozygous MuX mice after AAV CMV-Flpo treatment

(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder (L): GeneRuler 1kb Plus DNA Ladder(L); wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, magnification 11,2x (c) AAV genome persistence assay: detection of residual AAV CFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); loading control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (d) Summary of genotype and phenotype analysis.

Figure S6: Tissue genotyping and imaging of heterozygous MuX mice after AAV EF1a-Flpo treatment
(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder: GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, magnification 11,2x. (c) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); loading control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (d) Summary of genotype and phenotype analysis.
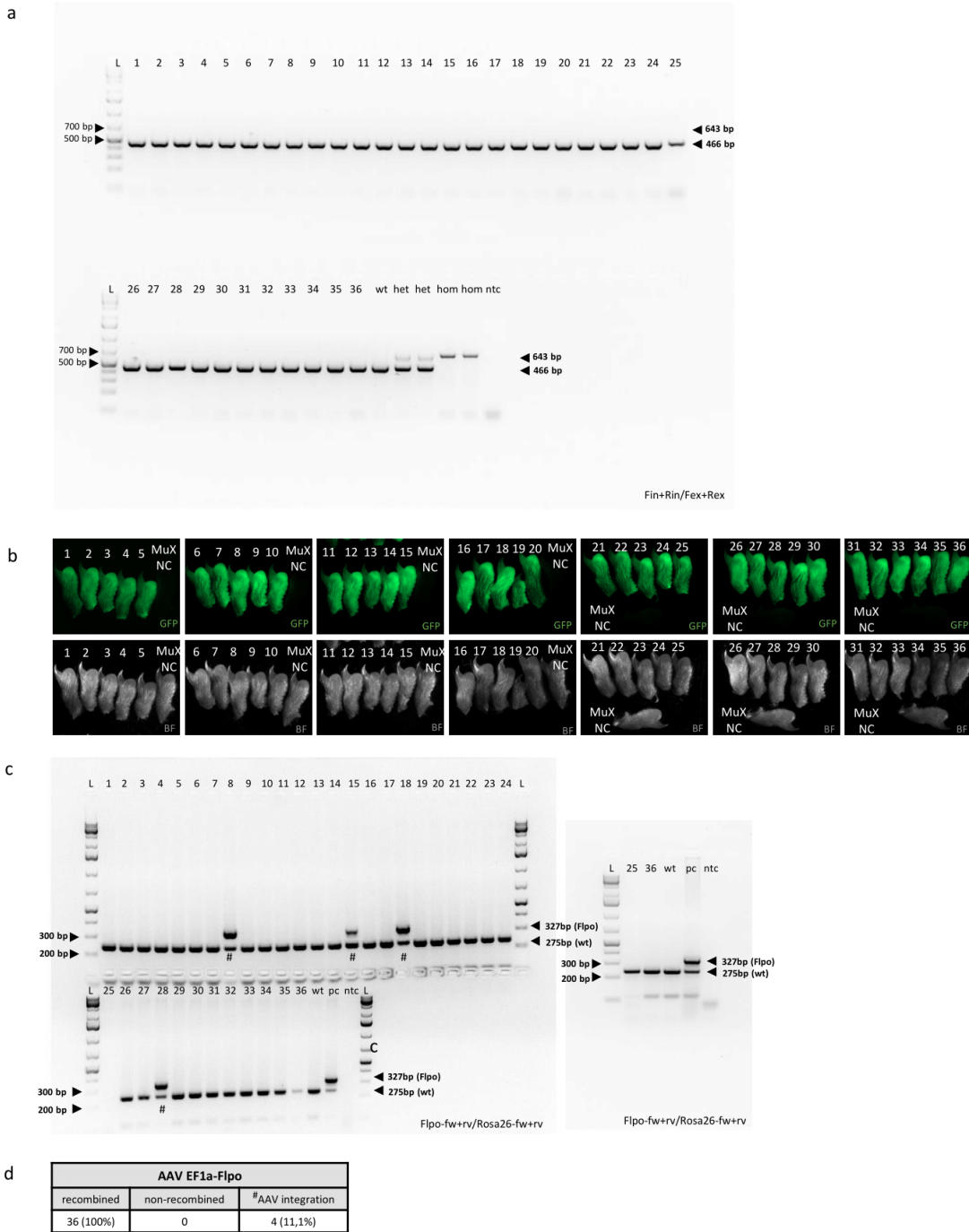
Figure S7: Tissue genotyping and imaging of heterozygous MuX mice after Cre protein electroporation

(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder (L): GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, magnification 11,2x. (C) Summary of genotype and phenotype analysis.

a

b

c

Dre protein	
recombined	non-recombined
13 (100%)	0
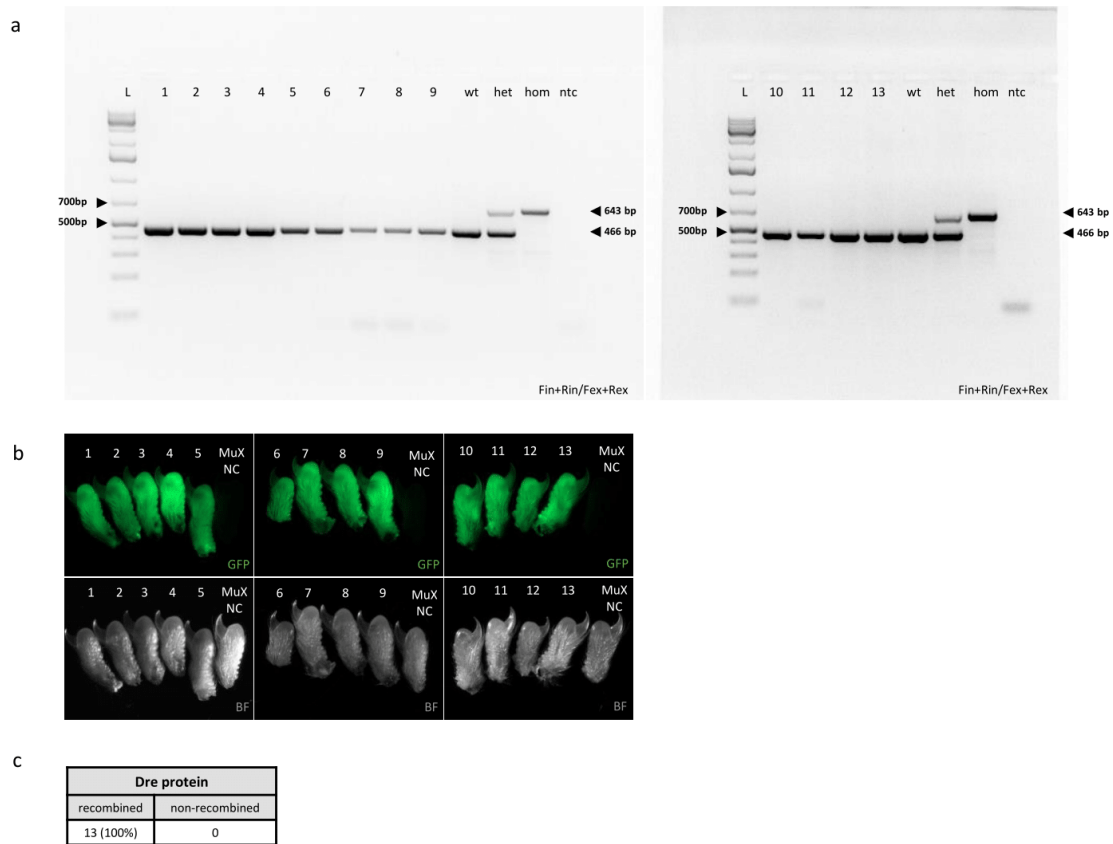
Figure S8: Tissue genotyping and imaging of heterozygous MuX mice after Dre protein electroporation

(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder (L): GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, magnification 11,2x. (c) Summary of genotype and phenotype analysis.

a

b

c

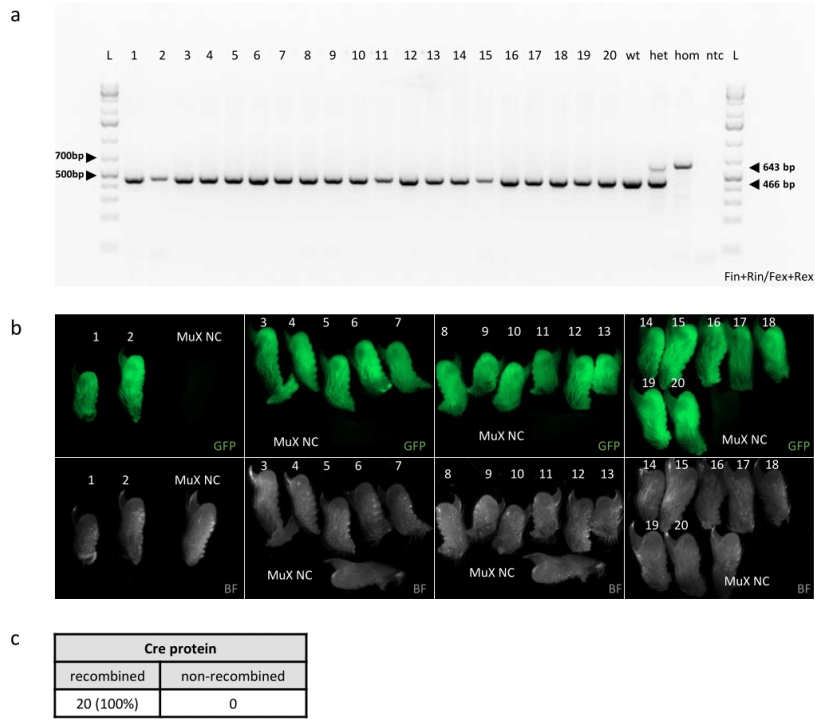| Vika mRNA	
recombined	**non-recombined
5 (71,4%)	2 (28,6%)
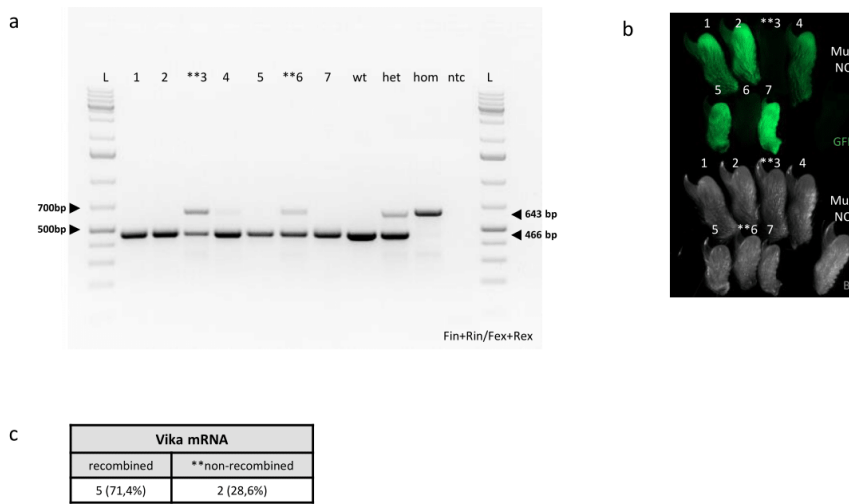
Figure S9: Tissue genotyping and imaging of heterozygous MuX mice after Vika mRNA microinjection

(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder (L): GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, magnification 11,2x. (c) Summary of genotype and phenotype analysis.
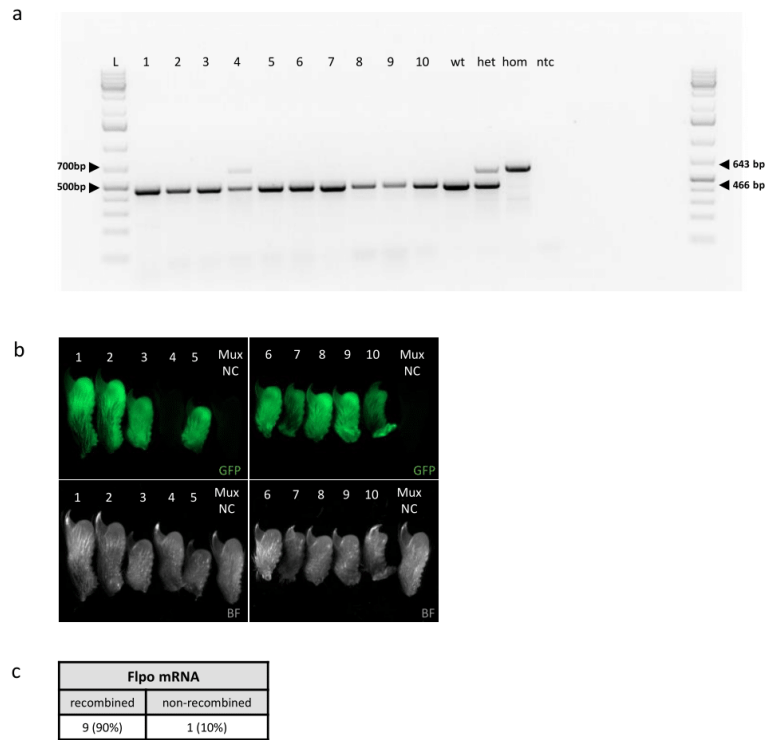
Figure S10: Tissue genotyping and imaging of heterozygous MuX mice after Flpo mRNA microinjection
(a) Mice were genotyped with Fin/Rin and Fex/Rex primers, ladder: GeneRuler 1kb Plus DNA Ladder; wt: wild-type tissue; het: heterozygote sample of mouse carrying a single MuX allele; hom: homozygote sample of mouse carrying two MuX alleles; Rosa26 locus without MuX cassette band corresponds to 466bp (Fex and Rex primers), 643bp band corresponds to non-recombined MuX cassette. (b) Clipped fingers were imaged for EGFP signal to confirm corresponding genotype, magnification 11,2x. (c) Summary of genotype and phenotype analysis
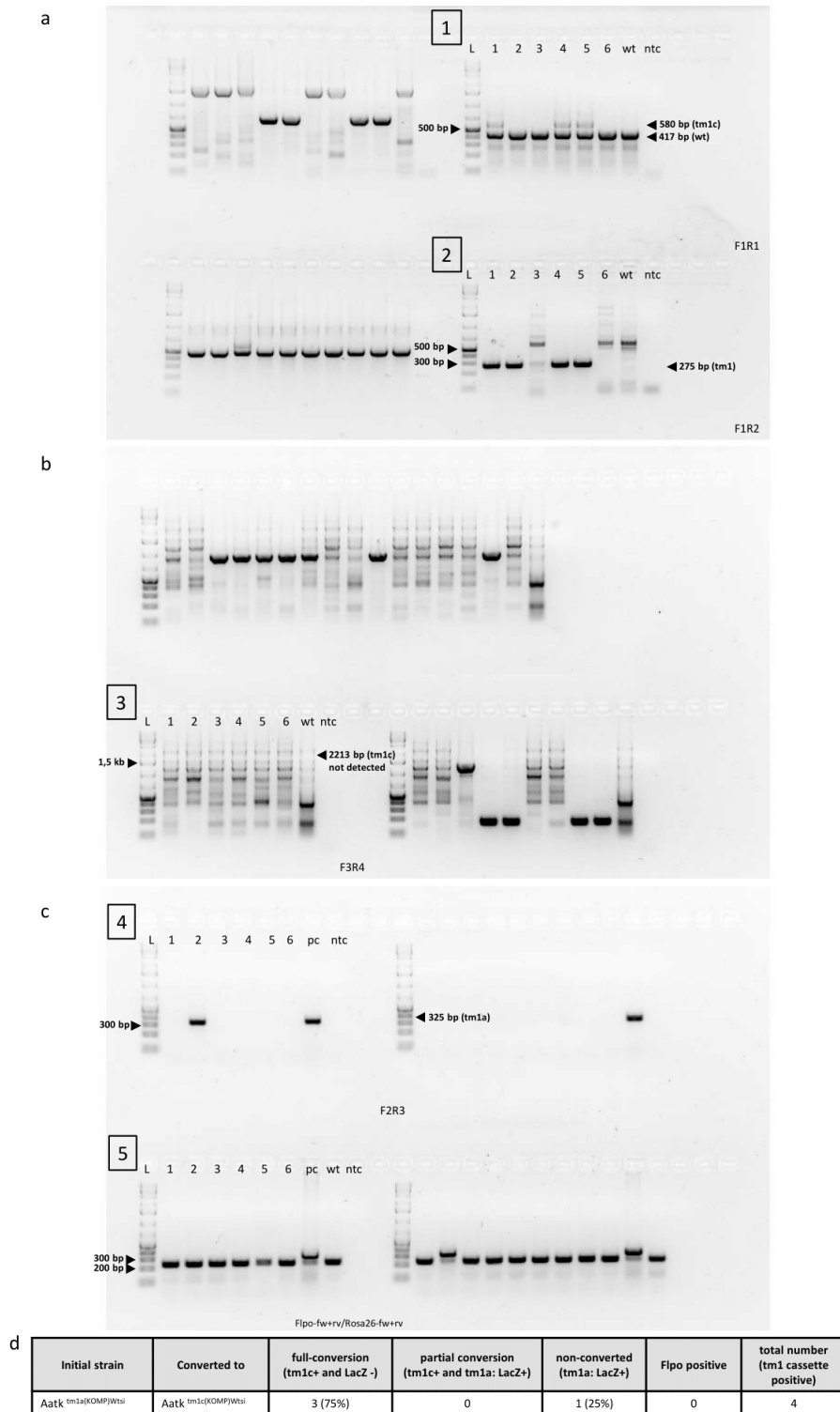
Figure S11: Detection of Aatk tm1c allele converted with AAV CMV-Flpo

(a-c) Detection of tm1a to tm1c conversion using multiple genotyping reactions (marked 1-5). All treated mice were born after in vitro fertilization of wild-type oocytes with heterozygote Aatk [tm1a(KOMP)Wtsi] sperm. (a1) F1R1 reaction detects presence of tm1c allele (580bp) and wild type allele (417bp). (a2) F1R2 reaction confirms presence of tm1 cassette (275bp). (b3) F3R4 reaction detects presence of tm1c (2213bp) allele. (c4) Detection of LacZ/tm1a cassette using F2R3 primers, positive band corresponds to 325bp. (c5) Integration assay: four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. Ladder (L): GeneRuler 1kb Plus DNA Ladder. (d) Summary of genotype analysis.

Initial strain	Converted to	full-conversion (tm1c+ and LacZ -)	partial conversion (tm1c+ and tm1a: LacZ+)	non-converted (tm1a: LacZ+)	Flpo positive	total number (tm1 cassette positive)
Aatk [tm1a(KOMP)Wtsi]	Aatk [tm1c(KOMP)Wtsi]	3 (75%)	0	1 (25%)	0	4

a

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 wt ntc

700 bp ▶
500 bp ▶

◀ 729 bp (tm1c)
◀ 610 bp (wt)

F1R1

b

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 wt ntc

300 bp ▶

◀ 259 bp (tm1)

F1R2

c

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 wt ntc

1 kb ▶

◀ 943 bp (tm1c)

F3R4

d

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 pc wt ntc

300 bp ▶

◀ 325 bp (tm1a)

F2R3

e

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 pc wt ntc

300 bp ▶
200 bp ▶

◀ 327 bp (Flpo)
◀ 275 bp (wt)

Flpo-fw+rv/Rosa26-fw+rv

f

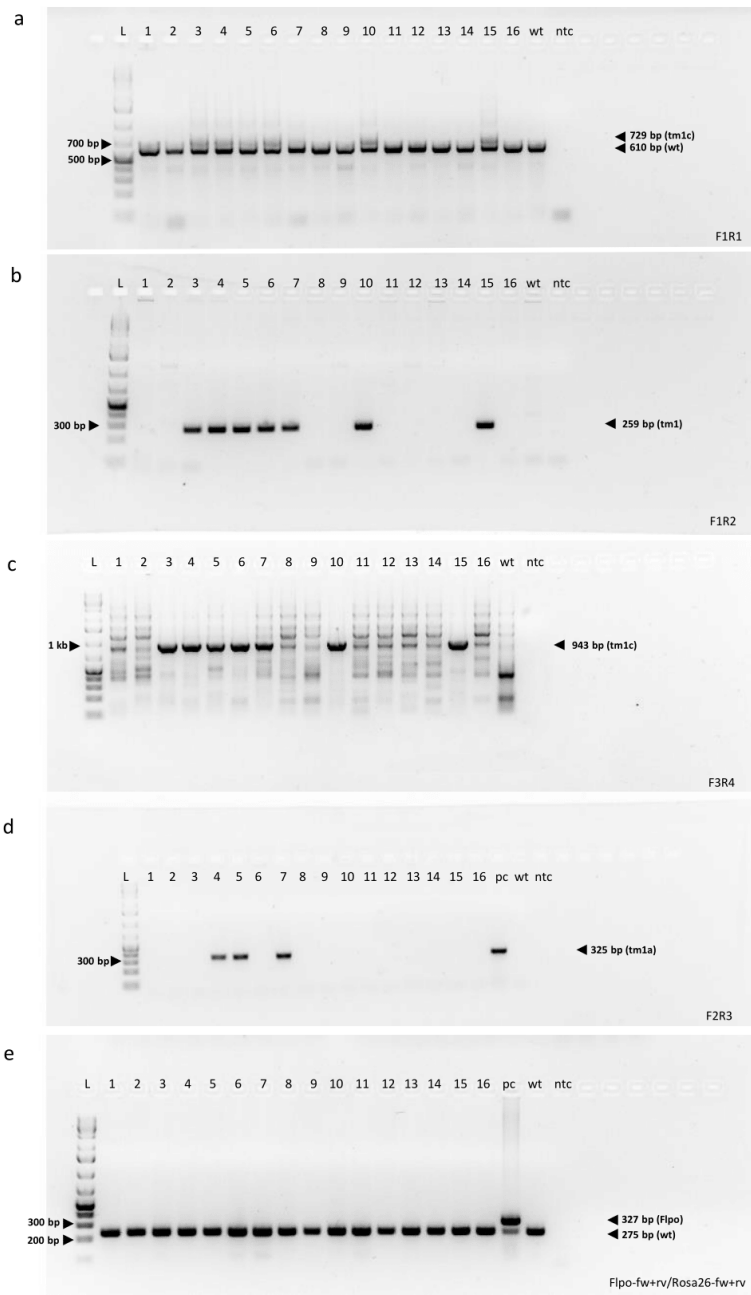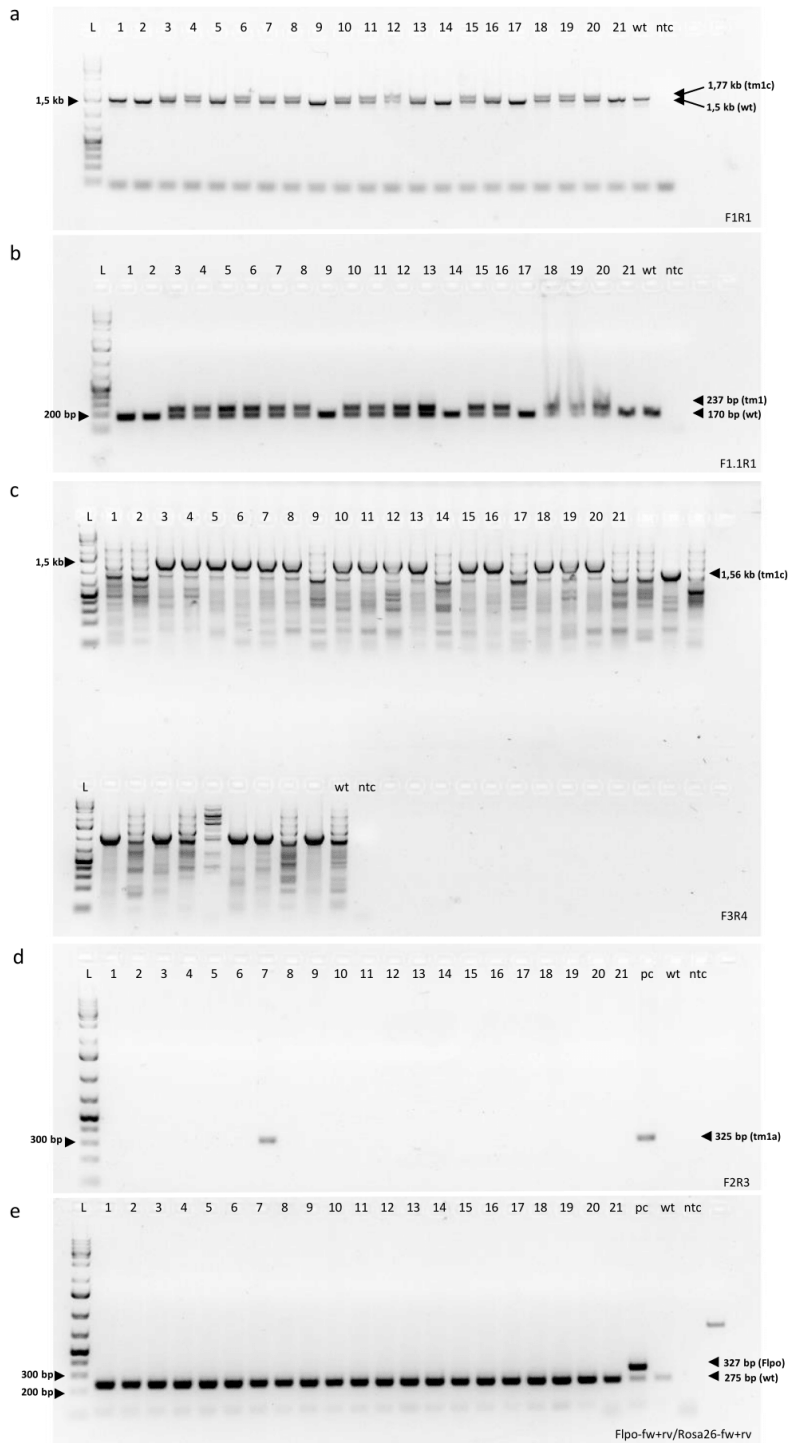| Initial strain	Converted to	full-conversion (tm1c+ and LacZ -)	partial conversion (tm1c+ and tm1a: LacZ+)	non-converted (tm1a: LacZ+)	Flpo positive	total number (tm1 cassette positive)
Cdh26 tm1a(KOMP)Wtsi	Cdh26 tm1c(KOMP)Wtsi	4 (57,1%)	3 (42,8%)	0	0	7

Figure S12: Detection of Cdh26 tm1c allele converted with AAV CMV-Flpo

Detection of tm1a to tm1c conversion using multiple genotyping reactions. All treated mice were born after in vitro fertilization of wild-type oocytes with heterozygote Cdh26tm1a(KOMP)Wtsi sperm. (a) F1R1 reaction detects presence of tm1c allele (729bp) and wild type allele (610bp). (b) F1R2 reaction confirms presence of tm1 cassette (259bp). (c) F3R4 reaction detects presence of tm1c (943bp) allele. (d) Detection of LacZ cassette using F2R3 primers, positive band corresponds to 325bp. (e) AAV genome persistence assay: detection of residual AAV CFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); loading control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (F) Summary of genotype analysis. Ladder (L): GeneRuler 1kb Plus DNA Ladder.

Figure S13: Detection of Atf2 tm1c allele converted with AAV EF1a-Flpo

Detection of tm1a to tm1c conversion using multiple genotyping reactions. All treated mice were born after in vitro fertilization of wild-type oocytes with heterozygote Atf2[tm1a(EUCOMM)Hmgu/Ph] sperm. (a) F1R1 reaction detects presence of tm1c allele (1,77kbp) and wild type allele (1,5kbp); (b) F1.1R1 reaction confirms presence of tm1 cassette (237bp), wild-type product corresponds to 170bp. (C) F3R4 reaction detects presence of tm1c (1,56kbp) allele. (d) Detection of LacZ/tm1a cassette using F2R3 primers, positive band corresponds to 325bp. (e) AAV genome persistence assay: detection of residual AAV CFlpo genome four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (f) Summary of genotype analysis. L (ladder): GeneRuler 1kb Plus DNA Ladder.

Initial strain	Converted to	full-conversion (tm1c+ and LacZ -)	partial conversion (tm1c+ and tm1a: LacZ+)	non-converted (tm1a: LacZ+)	Flpo positive	total number (tm1 cassette positive)
Atf2[tm1a(EUCOMM)Hmgu/Ph]	Atf2[tm1c(EUCOMM)Hmgu/Ph]	14 (93,3%)	1 (6,7%)	0	0	15

Figure S14: Detection of Ube2t tm1c allele converted with AAV EF1a-Flpo

Detection of tm1a to tm1c conversion using multiple genotyping reactions. All treated mice were born after in vitro fertilization of wild-type oocytes with heterozygote Ube2t[tm1a(EUCOMM)Hmgu>/WtsiPh] sperm. (a) F1R1 reaction detects presence of tm1c allele (752kbp) and wild type allele (665bp). (b) F1R2 reaction confirms presence of tm1 cassette (316bp). (c) F3R4 reaction detects presence of tm1c (884kbp) allele. (d) Detection of LacZ cassette using F2R3 primers, positive band corresponds to 325bp. (e) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (f) Summary of genotype analysis. L (ladder): GeneRuler 1kb Plus DNA Ladder.
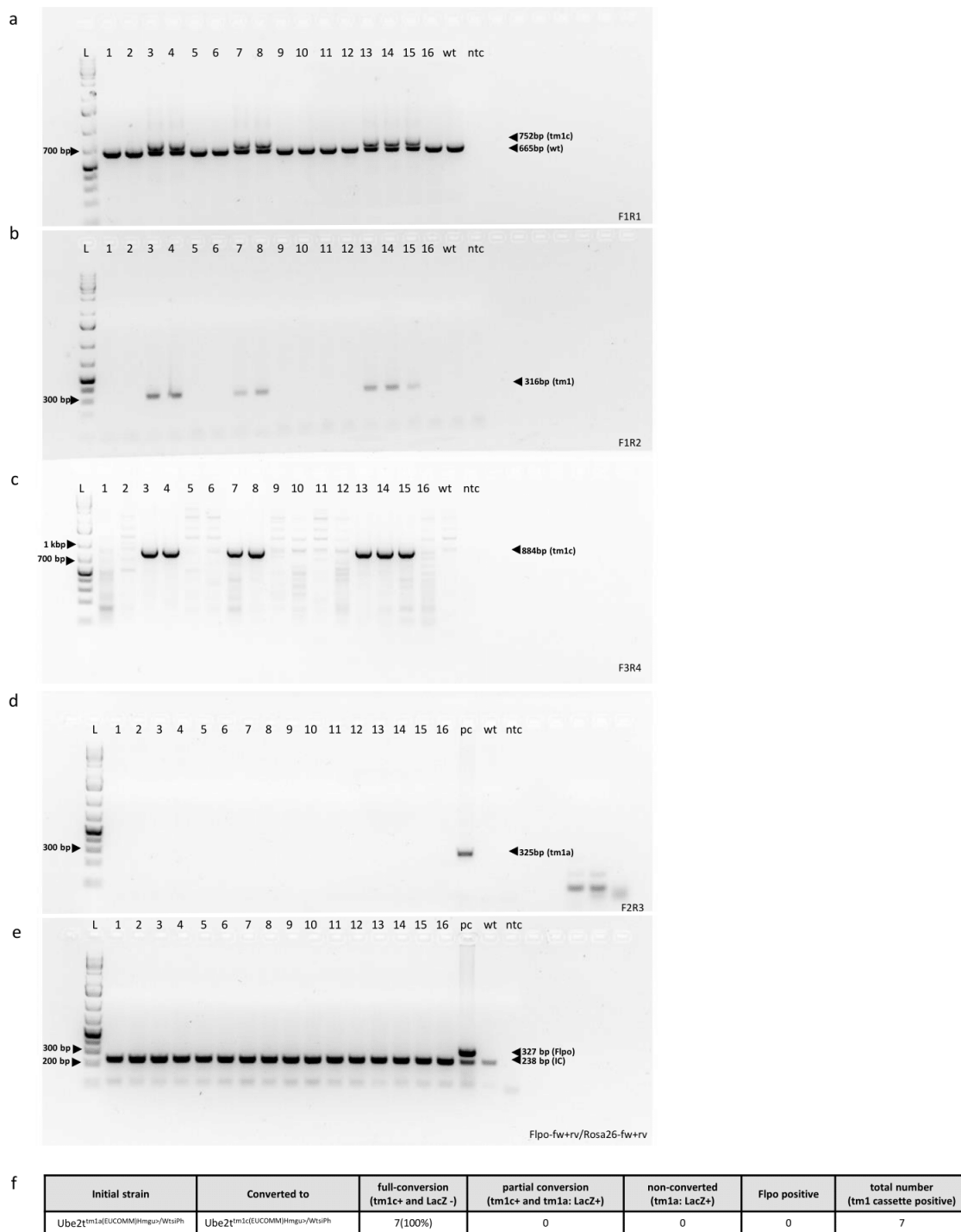
a

L 1 2 3 4 5 6 7 8 9 wt ntc

1.5 kb ▶

◀1652 bp (wt)

500 bp ▶

◀624 bp (tm1d)

F1R1

b

L 1 2 3 4 5 6 7 8 9 wt ntc

500 bp ▶

◀589bp (tm1c)
◀467bp (wt)

F1R2

c

L 1 2 3 4 5 6 7 8 9 wt ntc

1.5 kb ▶

◀1279bp (tm1c)

200 bp ▶

◀172bp (tm1d)

F3R4

d

L 1 2 3 4 5 6 7 8 9 wt ntc

300 bp ▶

◀325bp (tm1a)

F2R3

e

L 1 2 3 4 5 6 7 8 9 pc wt ntc

300 bp ▶
200 bp ▶

◀327bp (Flpo)
◀238bp (IC)

Flpo-fw+rv/Rosa26-fw+rv

f

L 1 2 3 4 5 6 7 8 9 wt pc pl ntc

1 kb ▶

◀951bp (Cre)

CMV-F+Cre_R

L 1 2 3 4 5 6 7 8 9 wt pc pl ntc

500 bp ▶

◀466bp (LC)

Fex+Rex

111

Initial strain	Converted to	full-conversion (tm1c+ and LacZ -)	partial conversion (tm1c+)	non-converted (tm1a: LacZ+)	Flpo positive	Cre positive	total number (tm1 cassette positive)
C4bp^{tm1a(KOMP)Wtsi/MbpMmucd}	C4bp^{tm1d(KOMP)Wtsi/MbpMmucd}	4 (80%)	1(20%)	0	1	0	5

Figure S15: Detection of C4bp tm1d allele converted with AAV EF1a-Flpo and AAV CMV-Cre

Detection of tm1a to tm1d conversion using multiple genotyping reactions. All treated mice were born after in vitro fertilization of wild-type oocytes with heterozygote C4bp^{tm1a(KOMP)Wtsi/MbpMmucd} . (a) F1R1 reaction detects tm1d allele (624bp) and wild-type (1652bp). (b) F1R2 reaction confirms presence of tm1c allele (589bp) and wild type allele (467bp). (c) F3R4 reaction detects presence of tm1c (1279bp) allele and tm1d (172bp). (d) Detection of LacZ/tm1a cassette using F2R3 primers, positive band corresponds to 325bp.(e) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (f) Detection of residual AAV Cre genome: CMV-F and Cre-R primers detect presence AAV CMV-Cre genome, band corresponding to 951bp; Fex and Rex primers produce 466bp band as a DNA loading control reaction; wt:wild-type sample, pc: wild-type DNA and 1ng/µ of AAV CMV-Cre plasmid, pl: AAV CMV-Cre plasmid, ntc: non-template. (g) Summary of genotype analysis. Ladder (L): GeneRuler 1kb Plus DNA Ladder.
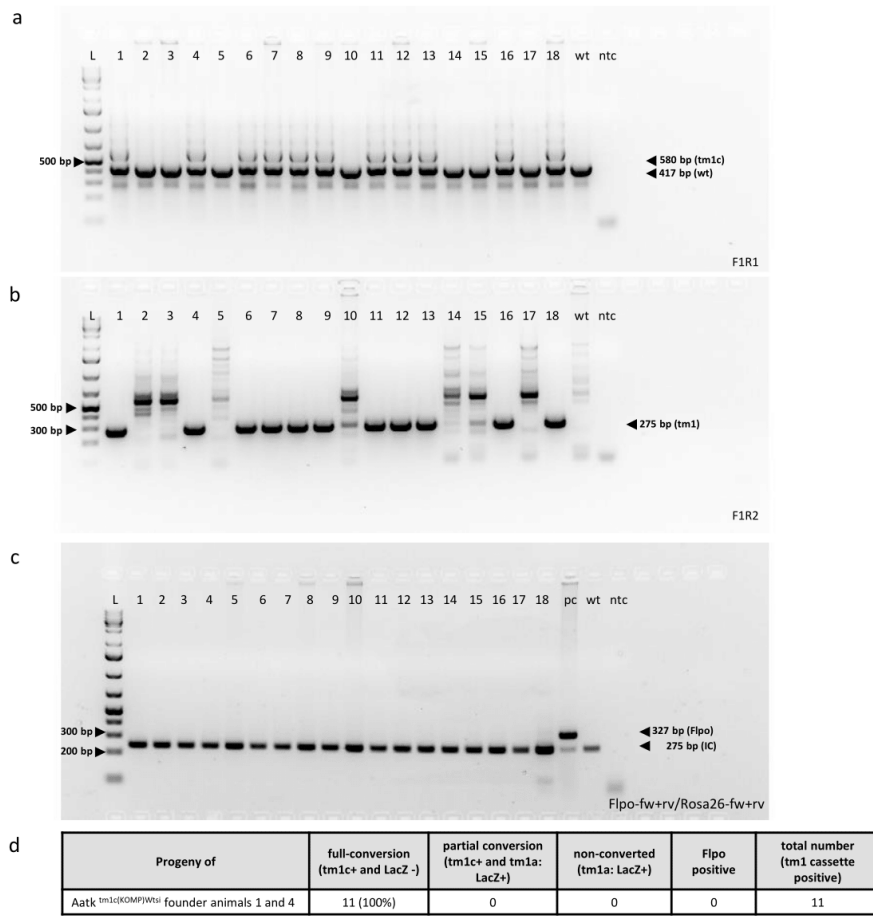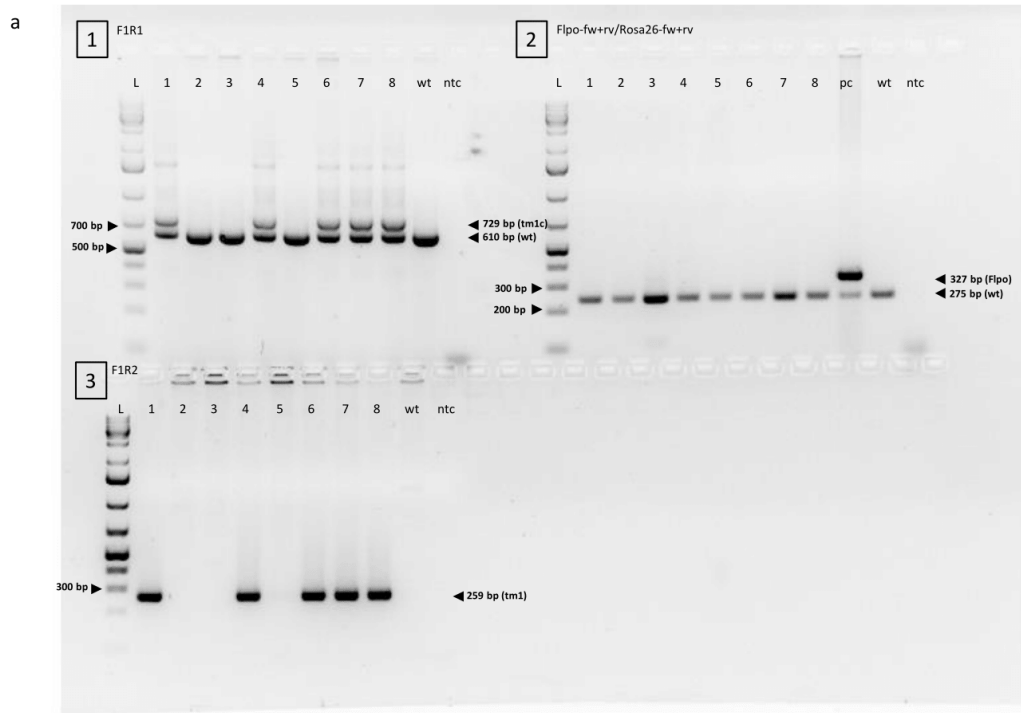
a

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 wt ntc

500 bp ◄ 580 bp (tm1c)
◄ 417 bp (wt)

F1R1

b

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 wt ntc

500 bp ◄
300 bp ◄ ◄ 275 bp (tm1)

F1R2

c

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 pc wt ntc

300 bp ◄ ◄ 327 bp (Flpo)
200 bp ◄ ◄ 275 bp (IC)

Flpo-fw+rv/Rosa26-fw+rv

d

Progeny of	full-conversion (tm1c+ and LacZ -)	partial conversion (tm1c+ and tm1a: LacZ+)	non-converted (tm1a: LacZ+)	Flpo positive	total number (tm1 cassette positive)
Aatk tm1c(KOMP)Wtsi founder animals 1 and 4	11 (100%)	0	0	0	11

Figure S16: Detection of Aatk tm1c allele in progeny of Flpo- converted animals

Confirmation of germ-line transmission of tm1c allele in progeny of AAV EF1a-Flpo converted Aatk tm1c animals. All analyzed mice were born from breeding of Aatk tm1c(KOMP)Wtsi animal 1 and 4 with wild-type (C57Bl/6NCr) animals. L (ladder): GeneRuler 1kb Plus DNA Ladder. (A) F1R1 reaction detects presence of tm1c allele (580bp) and wild type allele (417bp). (B) F1R2 reaction confirms presence of tm1 cassette (275bp). (C) AAV genome persistence assay: detection of residual AAV CFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (D) Summary of genotype analysis.

a

Panel 1: F1R1 — L, 1, 2, 3, 4, 5, 6, 7, 8, wt, ntc
700 bp ◄
500 bp ◄
◄ 729 bp (tm1c)
◄ 610 bp (wt)

Panel 2: Flpo-fw+rv/Rosa26-fw+rv — L, 1, 2, 3, 4, 5, 6, 7, 8, pc, wt, ntc
300 bp ◄
200 bp ◄
◄ 327 bp (Flpo)
◄ 275 bp (wt)

Panel 3: F1R2 — L, 1, 2, 3, 4, 5, 6, 7, 8, wt, ntc
300 bp ◄
◄ 259 bp (tm1)

b

Progeny of	full-conversion (tm1c+ and LacZ -)	partial conversion (tm1c+ and tm1a: LacZ+)	non-converted (tm1a: LacZ+)	Flpo positive	total number (tm1 cassette positive)
Cdh26 tm1c(KOMP)Wtsi founder animal 10	5 (100%)	0	0	0	5

Figure S17: Detection of Cdh26 tm1c allele in progeny of Flpo- converted animals

(a-b) Confirmation of germ-line transmission of tm1c allele in progeny of AAV EF1a-Flpo converted Cdh26 tm1d animals. All analyzed mice were born from breeding of Cdh26 tm1c(KOMP)Wtsi animal 10 with wild-type (C57Bl/6NCr) animals. L (ladder): GeneRuler 1kb Plus DNA Ladder. (a1) F1R1 reaction detects presence of tm1c allele (729bp) and wild type allele (610bp). (a2) AAV genome persistence assay: detection of residual AAV CFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); loading control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (a3) F1R2 reaction confirms presence of tm1 cassette (259bp (b) Summary of genotype analysis.
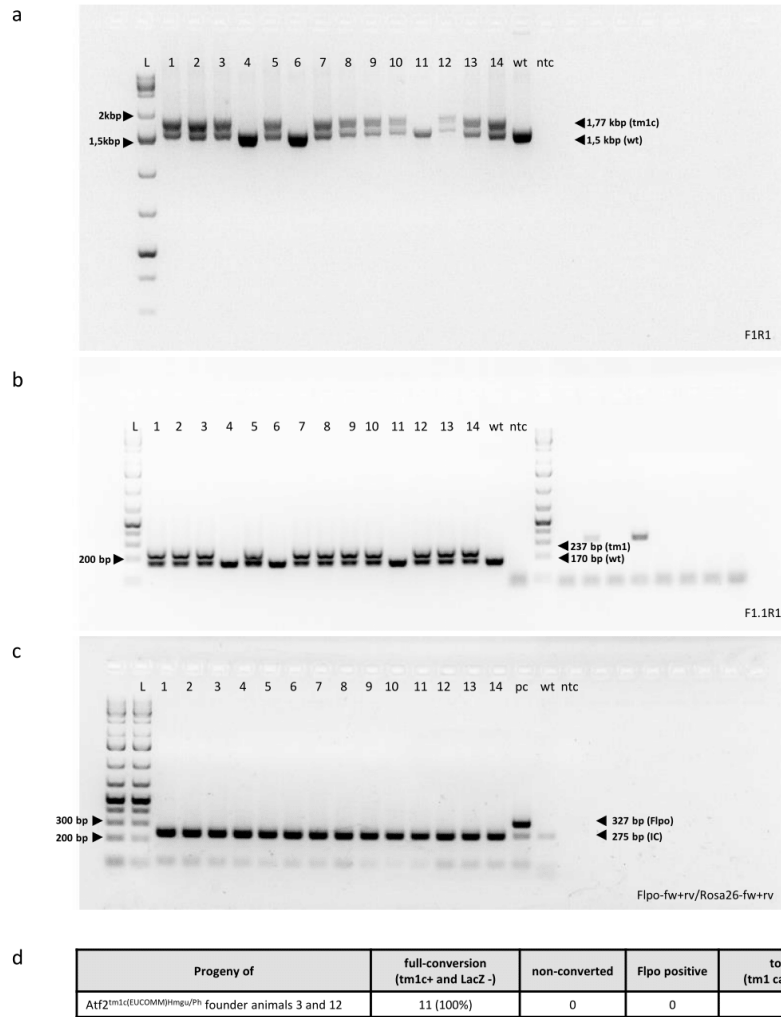
a

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt ntc

2kbp ►

1,5kbp ►

◄ 1,77 kbp (tm1c)
◄ 1,5 kbp (wt)

F1R1

b

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt ntc

200 bp ►

◄ 237 bp (tm1)
◄ 170 bp (wt)

F1.1R1

c

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 pc wt ntc

300 bp ►
200 bp ►

◄ 327 bp (Flpo)
◄ 275 bp (IC)

Flpo-fw+rv/Rosa26-fw+rv

d

Progeny of	full-conversion (tm1c+ and LacZ -)	non-converted	Flpo positive	total number (tm1 cassette positive)
Atf2tm1c(EUCOMM)Hmgu/Ph founder animals 3 and 12	11 (100%)	0	0	11

Figure S18: Detection of Atf2 tm1c allele in progeny of Flpo-converted animals

Confirmation of germ-line transmission of tm1c allele in progeny of AAV EF1a-Flpo converted Atf2 tm1c animals. All analyzed mice were born from breeding of Atf2tm1c(EUCOMM)Hmgu/Ph 3 and 12 with wild-type (C57Bl/6NCr) animals. (a) F1R1 reaction detects presence of tm1c allele (1,77kbp) and wild type allele (1,5kb). (b) F1.1R1 reaction confirms presence of tm1 cassette (237bp), wild-type product corresponds to 170bp. (c) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo: Flpo FR primers amplify AAV Flpo genome (327bp band); IC: internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (d) Summary of genotype analysis. Ladder (L): GeneRuler 1kb Plus DNA Ladder.
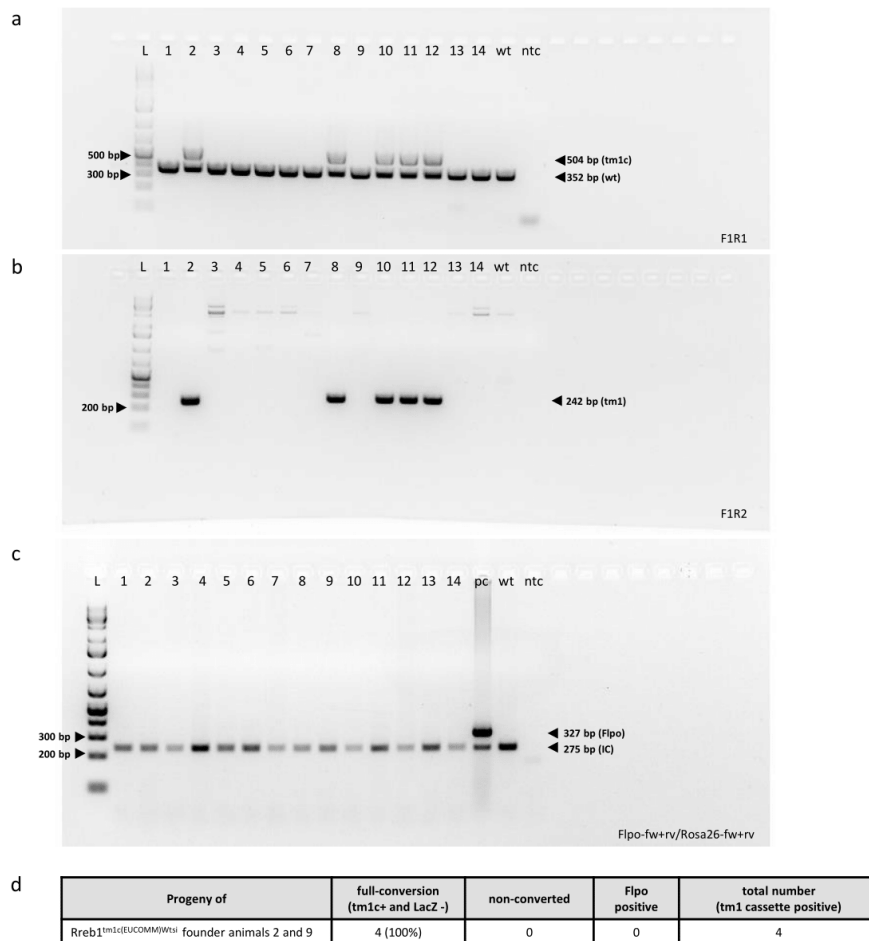
a

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt ntc

500 bp ▶
300 bp ▶

◀ 504 bp (tm1c)
◀ 352 bp (wt)

F1R1

b

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt ntc

200 bp ▶

◀ 242 bp (tm1)

F1R2

c

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 pc wt ntc

300 bp ▶
200 bp ▶

◀ 327 bp (Flpo)
◀ 275 bp (IC)

Flpo-fw+rv/Rosa26-fw+rv

d

Progeny of	full-conversion (tm1c+ and LacZ -)	non-converted	Flpo positive	total number (tm1 cassette positive)
Rreb1tm1c(EUCOMM)Wtsi founder animals 2 and 9	4 (100%)	0	0	4

Figure S19: Detection of Rreb1 tm1c allele in progeny of Flpo-converted animals

Confirmation of germ-line transmission of tm1c allele in progeny of AAV EF1a-Flpo converted Rreb1 tm1c animals. All analyzed mice were born from breeding Rreb1tm1c(EUCOMM)Wtsi 2 and 9 with wild-type (C57Bl/6NCr) animals. (a) F1R1 reaction detects presence of tm1c allele (504bp) and wild type allele (352bp). (b) F1R2 reaction confirms presence of tm1 cassette (242bp). (c) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo: Flpo FR primers amplify AAV Flpo genome (327bp band); IC: internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (d) Summary of genotype analysis. Ladder (L): GeneRuler 1kb Plus DNA Ladder.
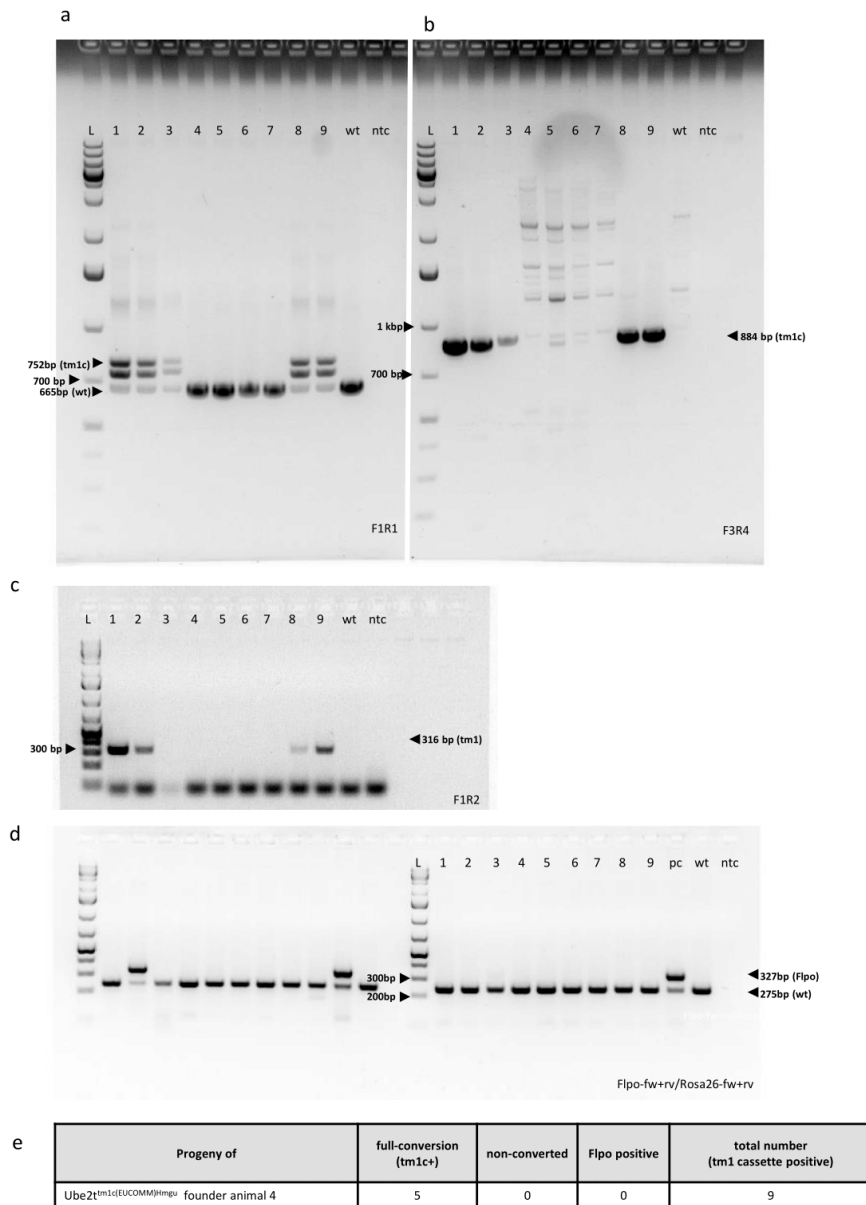
Figure S20: Detection of Ube2t tm1c allele in progeny of Flpo- converted animals

Confirmation of germ-line transmission of tm1c allele in progeny of AAV EF1a-Flpo converted Ube2t tm1c animals. All analyzed mice were born from breeding of Ube2t[tm1c(EUCOMM)Hmgu] 4 animal and wild-type (C57Bl/6NCr) animals. (a) F1R1 reaction detects presence of tm1c allele (752bp) and wild type allele (665bp). (b) F3R4 reaction detects presence of tm1c (884kbp) allele. (c) F1R2 reaction confirms presence of tm1 cassette (316bp). (d) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo: Flpo FR primers amplify AAV Flpo genome (327bp band); IC: internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (e) Summary of genotype analysis. Ladder (L): GeneRuler 1kb Plus DNA Ladder.
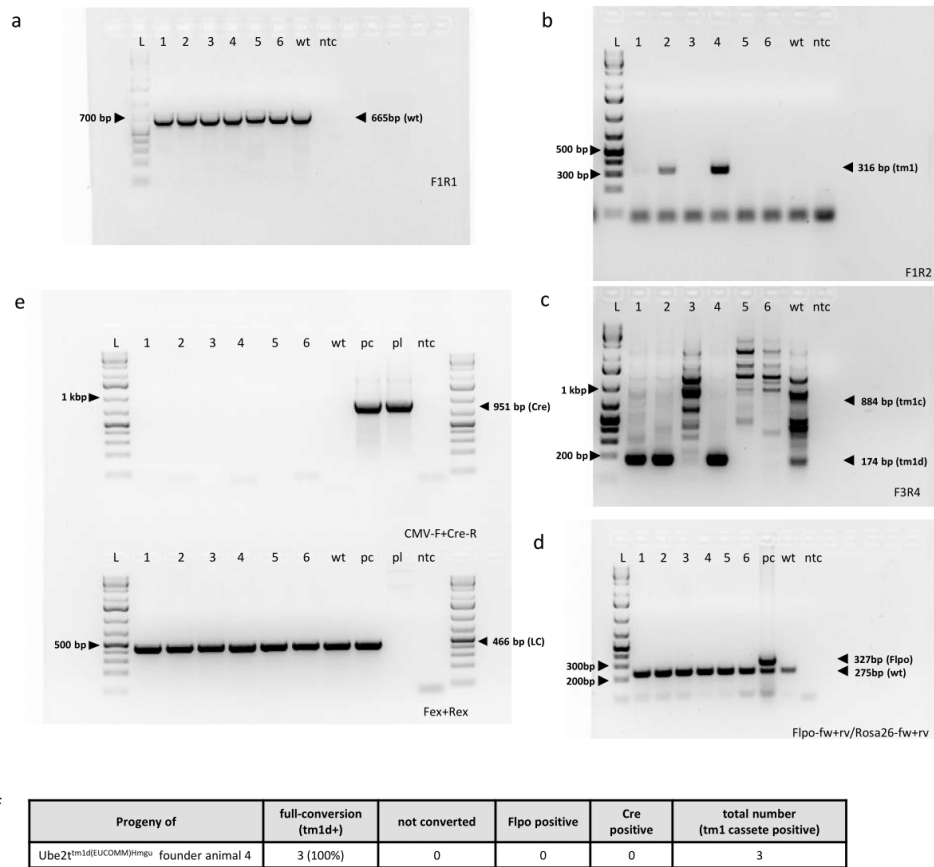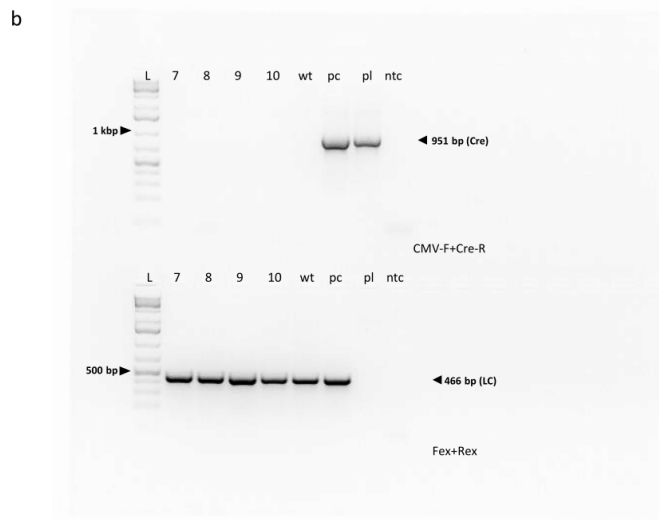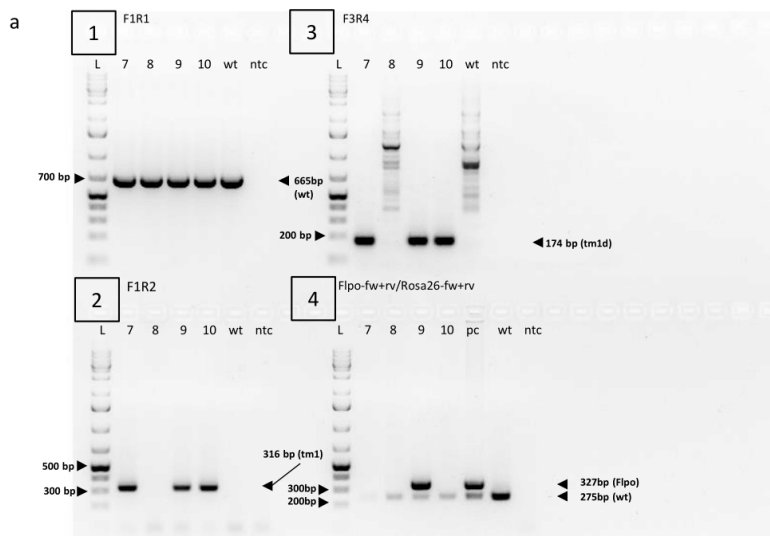
Figure S21: Detection of Ube2t tm1d allele in progeny of Flpo- and Cre- converted animals

Confirmation of germ-line transmission of tm1c allele in progeny of AAV EF1a-Flpo converted Ube2t tm1d animals. All analyzed mice were born from breeding of Ube2t[tm1d(EUCOMM)Hmgu] 4 wild-type (C57Bl/6NCr) animals. (a) F1R1 reaction detects presence of tm1c allele (752bp) and wild type allele (665bp). (b) F1R2 reaction confirms presence of tm1 cassette (316bp). (c) F3R4 reaction detects presence of tm1c allele (884bp) and wild type allele (174bp). (d) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo: Flpo FR primers amplify AAV Flpo genome (327bp band); IC: internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (e) Detection of residual AAV Cre genome, CMV-F and Cre-R primers detect presence AAV CMV-Cre genome, band corresponding to 951bp; Fex and Rex primers produce 466bp band as a DNA loading control reaction. (f) Summary of genotype analysis; wt:wild-type sample, pc: wild-type DNA and 1ng/μ of AAV CMV-Cre plasmid, pl: AAV CMV-Cre plasmid, ntc: non-template control. Ladder (L): GeneRuler 1kb Plus DNA Ladder.
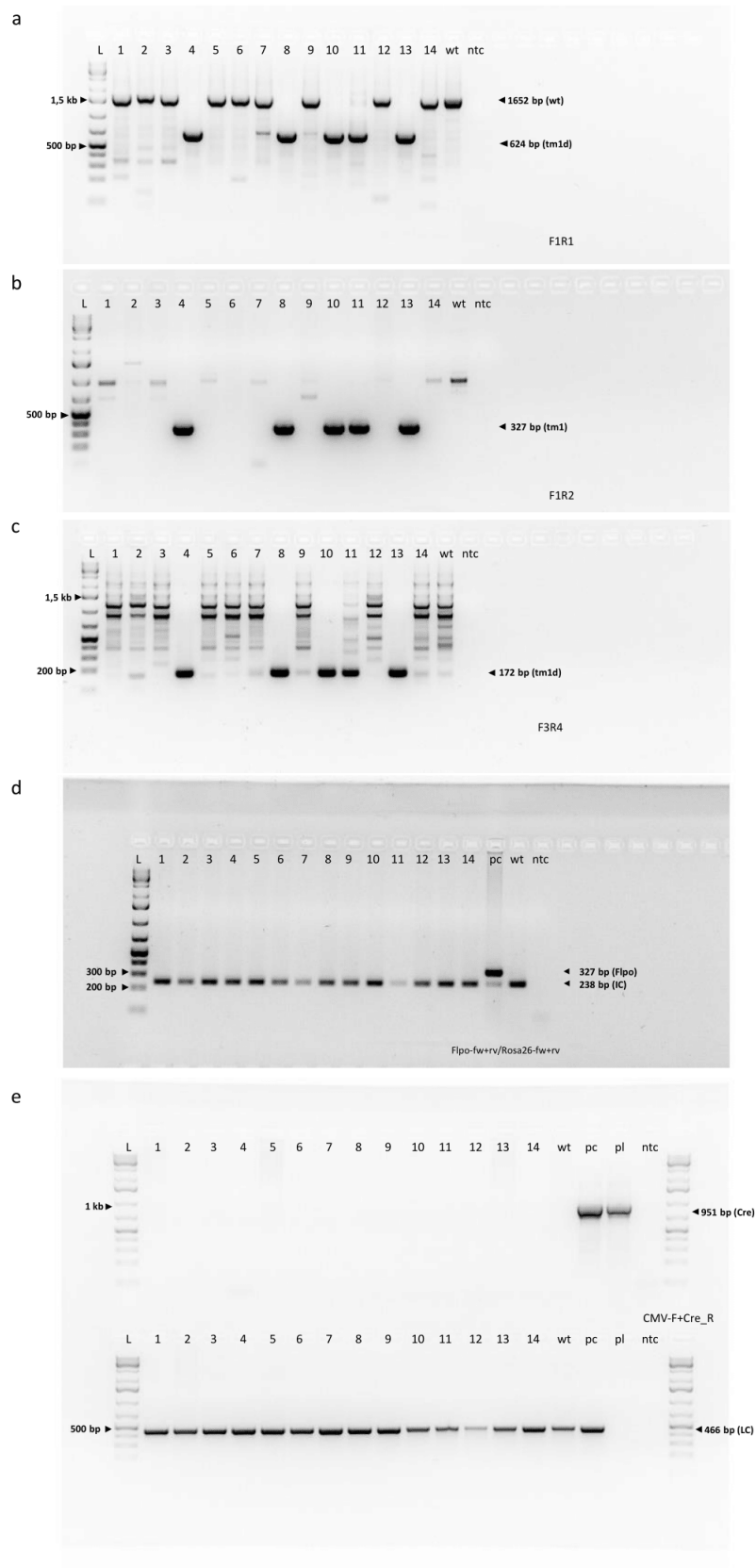
Figure S22: Detection of Ube2t tm1d allele in progeny of Flpo- and Cre- converted animals

Confirmation of germ-line transmission of tm1d allele in progeny of AAV EF1a-Flpo/CMV-Cre converted Ube2t tm1d animal. All analyzed mice were born from breeding of Ube2t[tm1d(EUCOMM)Hmgu] founder 2 with wild-type (C57Bl/6NCr). Ube2t tm1d founder 2 was tested positive for AAV Flpo genome. L (ladder): GeneRuler 1kb Plus DNA Ladder. (a1) F1R1 reaction detects presence of tm1c allele (752bp) and wild type allele (665bp). (a2) F1R2 reaction confirms presence of tm1 cassette (316bp). (a3) F3R4 reaction detects presence of tm1c allele (884bp) and wild type allele (174bp). (a4) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo: Flpo FR primers amplify AAV Flpo genome (327bp band); IC: internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (b) Detection of residual AAV Cre genome: CMV-F and Cre-R primers detect presence AAV CMV-Cre genome, band corresponding to 951bp; Fex and Rex primers produce 466bp band as a DNA loading control reaction; wt:wild-type sample, pc: wild-type DNA and 1ng/µ of AAV CMV-Cre plasmid, pl: AAV CMV-Cre plasmid, ntc: non-template control. (c) Summary of genotype analysis.

a

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt ntc

1,5 kb ▶

◀ 1652 bp (wt)

500 bp ▶

◀ 624 bp (tm1d)

F1R1

b

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt ntc

500 bp ▶

◀ 327 bp (tm1)

F1R2

c

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt ntc

1,5 kb ▶

200 bp ▶

◀ 172 bp (tm1d)

F3R4

d

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 pc wt ntc

300 bp ▶

◀ 327 bp (Flpo)

200 bp ▶

◀ 238 bp (IC)

Flpo-fw+rv/Rosa26-fw+rv

e

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt pc pl ntc

1 kb ▶

◀ 951 bp (Cre)

CMV-F+Cre_R

L 1 2 3 4 5 6 7 8 9 10 11 12 13 14 wt pc pl ntc

500 bp ▶

◀ 466 bp (LC)

120

f

Progeny of	full-conversion (tm1c+ and LacZ -)	partial conversion (tm1c+)	non-converted (tm1a: LacZ+)	Flpo positive	Cre positive	total number (tm1 cassette positive)
C4bp[tm1d(KOMP)Wtsi/MbpMmucd] founder animals 8 and 9	5 (100%)	0	0	0	0	5

Figure S23: Detection of C4bp tm1d allele in progeny of Flpo- and Cre- converted animals

Confirmation of germ-line transmission of tm1d allele in progeny of AAV EF1a-Flpo/CMV-Cre converted C4bp tm1d animals. All analyzed mice were born from breeding of C4bp[tm1a(KOMP)Wtsi/MbpMmucd] animals 8 and 9 with wild-type (C57Bl/6NCr) animals. (a) F1R1 reaction detects tm1d allele (624bp) and wild-type (1652bp). (b) F1R2 reaction confirms presence of tm1 allele (327bp). (c) F3R4 reaction detects presence of tm1c (1279bp) allele and tm1d (172bp). (d) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome (327bp band); internal control (Rosa26 FR primers) - 238bp band. wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (e) Detection of residual AAV Cre genome, CMV-F and Cre-R primers detect presence AAV CMV-Cre genome, band corresponding to 951bp; Fex and Rex primers produce 466bp band as a DNA loading control reaction; wt:wild-type sample, pc: wild-type DNA and 1ng/μ of AAV CMV-Cre plasmid, pl: AAV CMV-Cre plasmid, ntc: non-template. (f) Summary of genotype analysis. Ladder (L): GeneRuler 1kb Plus DNA Ladder.
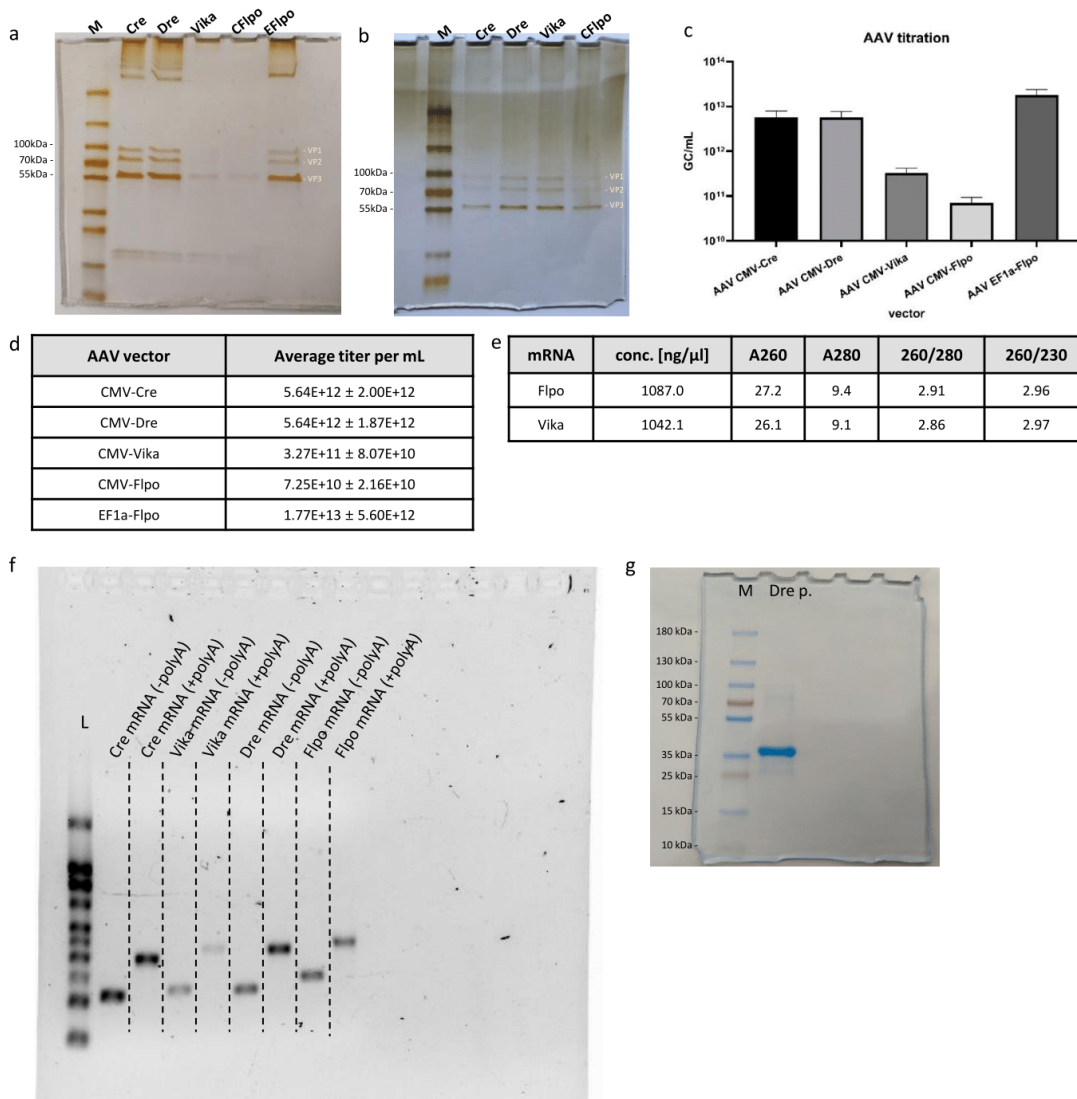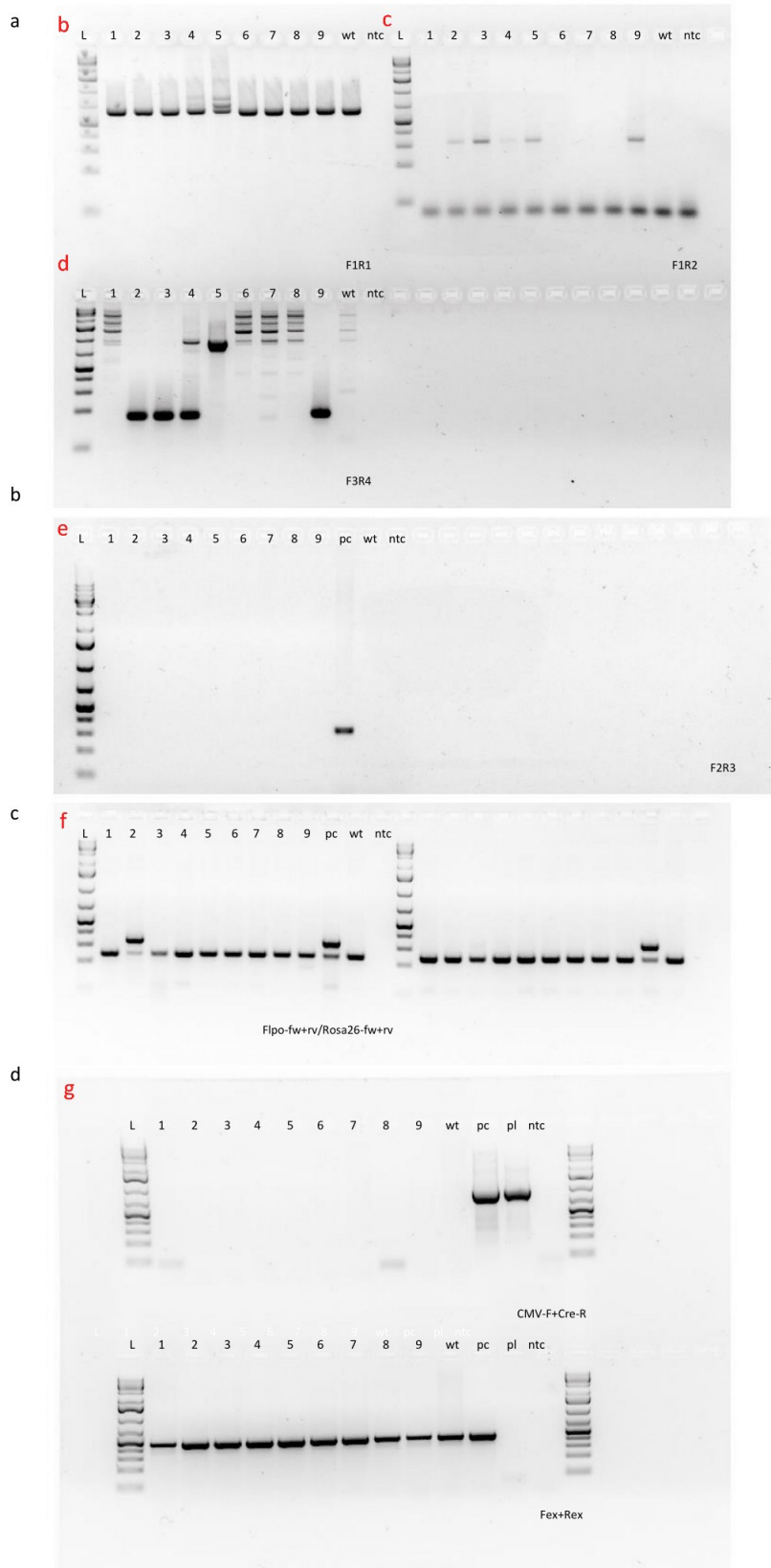
Fig.S24 Molecules and vectors quality control

(a) Silver-stained SDS-PAGE gel of AAV encoding Cre, Dre, Vika and Flpo recombinases, M - PageRuler™ Plus Prestained Protein Ladder (26620, Thermo Scientific). Separation of equal volume (9µl) of each AAV vector per lane. CFlpo – AAV CMV-Flpo. EFlpo- AAV EF1a-Flpo is commercially supplied vector and serves as purity control in the analysis. (b) Separation of equal titer of AAV vectors per lane. CFlpo – AAV CMV-Flpo. AAV viral proteins: VP1, VP2, VP3. (c) Graph depicting average titer of corresponding vectors. (d) Table showing titers of individual AAV vectors. (e) Spectral analysis of synthetized mRNA used for microinjection. (f) Assessment of integrity of SSR mRNA before and after polyA tailing. Cre and Dre mRNA were used as synthesis controls L (ladder): Millennium™ RNA Marker. (g) Purity assesment of synthetic Dre protein (41,7kDa) using Coomassie Blue staining, M - PageRuler™ Plus Prestained Protein Ladder (26620). Thermo Scientific), Dre p.- purified Dre protein (used for electroporation in zygotes).

a

b L 1 2 3 4 5 6 7 8 9 wt ntc   c L 1 2 3 4 5 6 7 8 9 wt ntc

F1R1   F1R2

d L 1 2 3 4 5 6 7 8 9 wt ntc

F3R4

b

e L 1 2 3 4 5 6 7 8 9 pc wt ntc

F2R3

c

f L 1 2 3 4 5 6 7 8 9 pc wt ntc

Flpo-fw+rv/Rosa26-fw+rv

d

g L 1 2 3 4 5 6 7 8 9 wt pc pl ntc

CMV-F+Cre-R

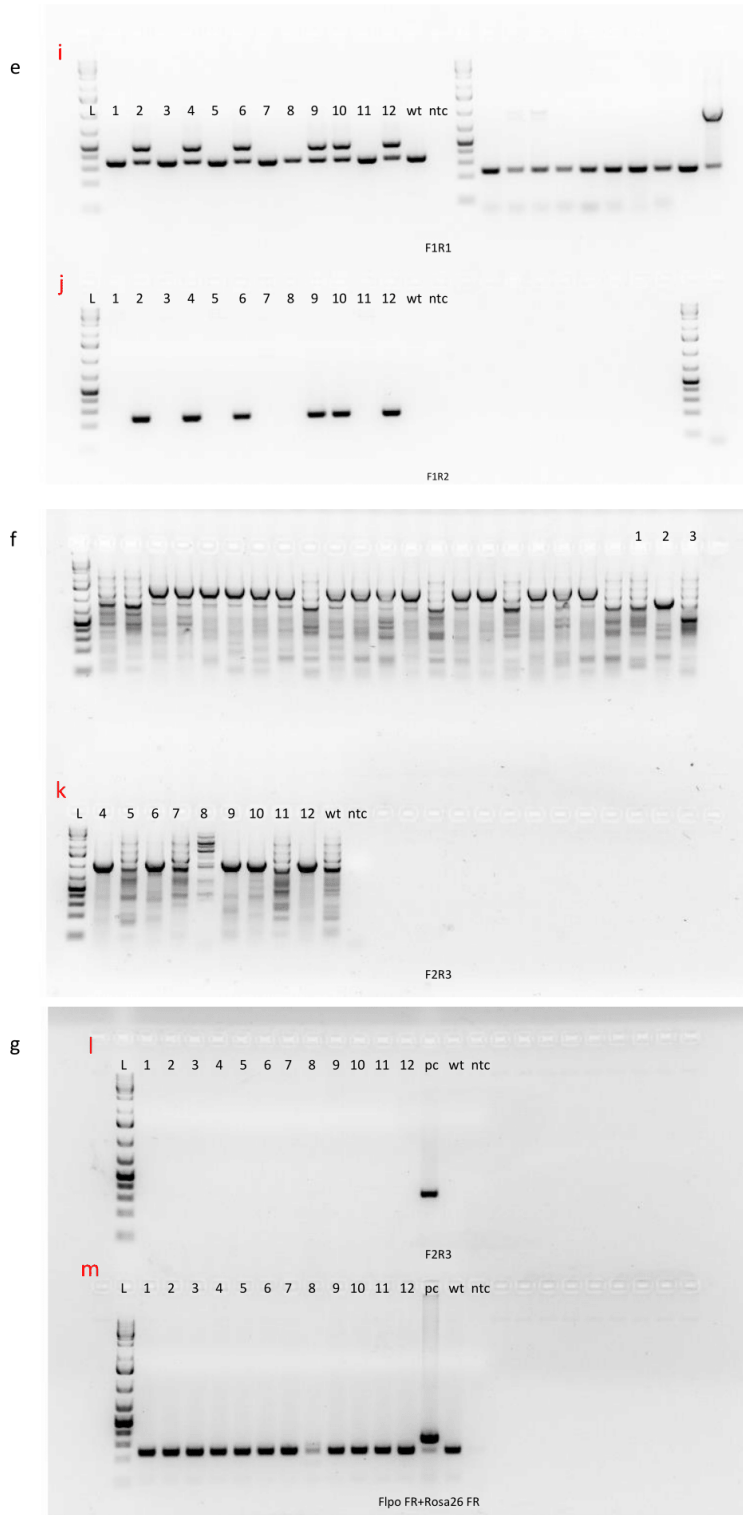L 1 2 3 4 5 6 7 8 9 wt pc pl ntc

Fex+Rex

Figure S25: EUCOMM/KOMP allele conversion using AAV Flpo and/or AAV Cre vectors (uncropped gels of Fig.3)

Images a-g represent the uncropped gel images corresponding to gels in Fig. 3 (red letters b-m). (red b-e) Ube2t^{tm1d (EUCOMM)Hmgu} genotyping with corresponding primers to detect tm1d allele. (red f and m) AAV genome persistence assay: detection of residual AAV EFlpo genome, four primers reaction, Flpo FR primers amplify AAV Flpo genome; internal control (Rosa26 FR primers). wt: C57Bl/6NCrl genomic DNA, pc: positive control - Gt(Rosa)26Sor(CAG-Flpo,-EYFP) genomic DNA. (red g) Detection of residual AAV Cre genome, CMV-F and Cre-R primers detect presence AAV CMV-Cre genome; Fex and Rex primers amplify a DNA loading control reaction; wt:wild-type sample, pc: wild-type DNA and 1ng/μ of AAV CMV-Cre plasmid, pl: AAV CMV-Cre plasmid, ntc: non-template control. (red h) Ube2t^{tm1d (EUCOMM)Hmgu} genotyping summary. (red i-l) Rreb1^{tm1c (EUCOMMWtsi} genotyping with corresponding primers to detect tm1c allele. L(ladder): GeneRuler 1kb Plus DNA.
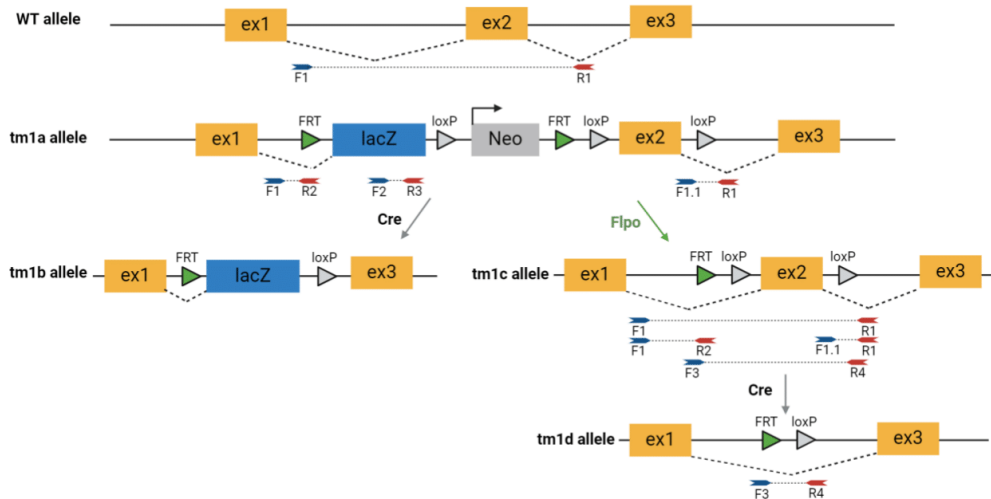
124

Supplementary Method:

Figure S26: Genotyping Strategies for Functional Analysis of Target Genes in EUCOMM/KOMP Alleles

The knock-out first EUCOMM/KOMP allele, tm1a, allows for the analysis of target gene expression based on the LacZ reporter and facilitates selection based on neomycin resistance. The tm1b allele is result of Cre recombination. tm1b enables expression analysis based on LacZ expression and disrupts gene function by deleting a critical exon. tm1c is result of Flpo conversion which gives rise to conditional form of the gene of interest, flanking a critical exon with two loxP sites. tm1d arises from the converted tm1c allele, leading to gene knock-out without reporter or selection properties. The F1 (or F1.1) and R1 primers detect the wild-type allele and tm1c allele, while the F1 and R2 primers detect the presence of the tm1(a, b, c, d) cassette in the genome. The F2 and R3 primers detect the LacZ cassette, and the F3 and R4 primers detect either the tm1c allele, tm1d, or both alleles.

Aatk^{tm1a(KOMP)Wtsi} Flpo-converted mice were genotyped for presence of tm1 allele (product size: 275 bp) with Aatk F1 and tm1-R2 primers. Further, Aatk F1 and R1 primers were used to detect wt (417 bp) and tm1c (580 bp) alleles. *LacZ* (product size: 325 bp) specific primers LacZ-F2 and LacZ-R3 were used to confirm complete conversion to tm1c allele by removal of *LacZ* cassette.

Cdh26^{tm1a(KOMP)Wtsi} Flpo-converted mice were genotyped for presence of tm1 allele (product size: 259 bp) with Cdh26 F1 and tm1-R2 primers. Further Cdh26 F1 and R1 primers were used to detect wt (610 bp) and tm1c (729 bp) alleles. *LacZ* (product size: 325 bp) specific primers LacZ-F2 and LacZ-R3 were used to confirm complete conversion to tm1c allele by removal of *LacZ* cassette.

Ube2t^{tm1a(EUCOMM)Hmgu} converted mice were genotyped for presence of tm1 allele with Ube2t F1 and tm1-R2 primers (product size: 318 bp). Further Ube2t F1 and R1 primers were uses to detect wt (product size: 665 bp) and tm1c (product size: 752 bp) alleles. LacZ-F2 and LacZ-R3 primers (product size: 325 bp) were used to confirm complete conversion to tm1c allele and absence of *LacZ* cassette. Conversion to tm1d allele was confirmed with tm1c-F3 and floxed-R4 primers amplifying a specific product of 174 bp. In case of tm1c allele, tm1c-F3/ floxed-R4 primers generate 884bp product.

Atf2^{tm1a(EUCOMM)Hmgu/Ph} Flpo-converted mice were genotyped for presence of tm1 allele with Atf2-F1.1 and Atf2-R1 primers which amplify wild-type allele giving 170bp product, and in presence of tm1 cassette, product of 237bp. Atf2-F1 and Atf2-R1 primers (product of 1721 bp) show presence of tm1c allele, where wt allele gives product of 1506 bp. LacZ-F2 and LacZ-R3 primers were used to confirm complete conversion to tm1c allele and absence of *LacZ* cassette. PCR reaction with tm1c-F3 and floxed-R4 primers, aligning upstream to FRT and downstream to *loxP*, confirms conversion and integrity tm1c allele with 1563bp product.

Rreb1^{tm1a(EUCOMM)Wtsi} Flpo-converted mice were genotyped for presence of tm1c allele with Rreb1-F1 and Rreb1-R1 primers which amplify wild-type allele giving 352bp product, and in presence of tm1c allele 504 bp product. To distinguish mice with tm1 allele and wt, primers Rreb1-F1 and tm1-R2 were used. tm1 positive mice give 242 bp product. *LacZ* (325 bp) specific primers, LacZ-F2 and LacZ-R3, were used to confirm complete conversion to tm1c allele and absence of *LacZ* cassette. PCR reaction with tm1c-F3 and floxed-R4 primers, aligning upstream to FRT and downstream to *loxP*, confirms conversion and integrity tm1c allele with 1047bp product.

C4bp^{tm1a(KOMP)Wtsi/MbpMmucd} Flpo-converted mice were genotyped for presence of tm1 allele with C4bp F1 and R2 primers (wt product size: 467 bp, tm1c product size: 589 bp, tm1d: no product). Further C4bp F1 and R1 primers were used to detect wt (product size: 1652 bp), tm1c (product size: 1785 bp) and tm1d (product size: 678 bp) alleles. LacZ-F2 and LacZ-R3 primers (product size: 325 bp) were used to confirm complete conversion to tm1c allele and absence of *LacZ* cassette. Conversion to tm1d allele was confirmed with tm1c-F3 and floxed-R4 primers amplifying a specific product of 172 bp. In case of tm1c allele, tm1c-F3/floxed-R4 primers generate 1279bp product.

Despite sequential differences, all listed primers correspond to the depicted primers in Fig.S21. The approximate site of binding is indicated based on their number (e.g., F1, R3, etc.). Their correspondence to specific genes is given by their names (e.g., Ube2t-F1, Rreb1-R1, etc.). Primers such as LacZ-F2, LacZ-R3, tm1-R2, floxed-R4, and tm1c-F3 are common for all used strains, and can be used for any tm1 allele, regardless of the gene of interest.

All AAV SSR treated mice were analyzed for potential persistance of AAV genome in integrated or episomal state. Forward primer binding CMV promoter (CMV-Fw) was used to detect potentially active persisting AAV genomes for AAV CMV-Cre, Dre and Vika vectors. Reverse primer was used to detect a specific AAV SSR genome based on treated group identity (Cre-Rv, Dre-Rv, Vika-Rv). Fex and Rex primers were used for amplification of wt Rosa26 locus as loading control reaction. Flpo-fw and rv primers were used in four primers reaction (with Rosa26 Fw and Rv primers) to detect *Flpo* persistence in cells. Rosa Fw and Rv primers align with Rosa26 locus and their product serves as internal control.

Primers and oligos:

Cloning oligos	sequence (5'>3')
CDS Dre fw	agaaggagatataccatgcctaagaagaagaggaagg
CDS Dre rv	atggtggtggtgatggctatccatcagtcgagaattgg
CDS Vika fw	agaaggagatataccatgaagaaaaagcggaaagtgacc
CDS Vika rv	atggtggtggtgatgcaaagagaattcctcgactctagac
CDS Flpo fw	agaaggagatataccatggctcctaagaagaagaggaa
CDS Flpo rv	atggtggtggtgatggatccgcctgttgatgtagc
AAV CDS Cre fw	ttcgaacatcgattgagaattctgcagtcgacggta
AAV CDS Cre rv	gcaggtcgactctagaatggctaatcgccatcttcc
AAV CDS Dre fw	aacatcgattgaattctgaatgcctaagaagaagaggaagg
AAV CDS Dre rv	gagcagcgctgctcgaggcacgttcaaatttcgcagcagc
AAV CDS Vika fw	ttcgaacatcgattgaattccccgccaccatgcctaa
AAV CDS Vika rv	gagcagcgctgctcgaggcacgagctctgactcacaaagag
AAV CDS Flpo fw	ttcgaacatcgattgaattcgcagtcgacggtaccggat
AAV CDS Flpo rv	gagcagcgctgctcgaggcatgcacctgaggagtgcggtta

AAV titration primers	sequence (5'>3')
AAV ITR fw	GGAACCCCTAGTGATGGAGTT
AAV ITR rv	CGGCCTCAGTGAGCGA

Genotyping primers	sequence (5'>3')
Fex (MuX)	gagccataactgcagactt
Rex (MuX)	cggatttagccacatccata
Fin (MuX)	cctctgctaaccatgttcat
Fin (MuX)	ttaccaatatccaggccaac

Name	primer	Alternative name	sequence (5'>3')
common primers	F2	LacZ-internal-Fw	gcgaatacctgttccgtcat
	R3	LacZ-internal-Rv	caaaaatccatttcgctggt
	R2	CAS_R1_Term	tcgtggtatcgttatgcgcc
	F3	tm1c_F	aaggcgcataacgataccac
	R4	Floxed LR	actgatggcgagctcagacc
Aatk	F1	Aatk_41505_F	GGTGTAAGTCTTGGGGGCAC
	R1	Aatk_41505_R	GTTTGGATGAGGCAGGAAGC
Cdh26	F1	Cdh26_35701_F	TTTGTCCCCTTGACATTCCC
	R1	Cdh26_35701_R	GATGAGCAGTGCAGTCCCAG
Ube2t	F1	Ube2t_278300_F	catggtaaaaggcgcaaacc
	R1	Ube2t_278300_R	ccccaagttcagcaggaaac
Atf2	F1	5'Atf2fw	gagatatcgttgaaatagttagg
	F1.1	3'Atf2Fw	gaaaaggtgttatttgaagatatg
	R1	3'Atf2Rv	catcatgatgtaacttaacacttaac
Rreb1	F1	Rreb1_81122_F	aagcccaaagtcctcgtgtc
	R1	Rreb1_81122_R	gcagtccctcctccagtttg
C4bp	F1	CSD-C4bp_F	gactgaaggggtcatgcaaagaagc
	R1	CSD-C4bp_R	tgatgtgctcactcccagacaaagg
	R2	CSD-C4bp-ttR	tctgaacagacaccatgaccaaggg

Integration assay primers	sequence (5'>3')
CMV-Fw	ctttccacgtcctggtgtct
Cre-Rv	tgcgaacctcatcactcgtt
Dre-Rv	ccatctggcccagctatgaa
Vika-Rv	ctttccacgtcctggtgtct
Flpo-Fw	ttttcagcgccagaggcaggatcg
Flpo-Rv	gtcctggccacggcggaggc
Rosa26-Fw	aaagtcgctctgagttgttat
Rosa26-Rv	cctttaagcctgcccagaag

4.4. Part 3: rAAV Purification through the strategic utilization of EVs (unpublished)

In order to make rAAV vectors more accessible for our research focused on transgenesis and gene delivery, we developed a scalable method for rAAV production and purification. The original aim of using extracellular vesicle associated rAAV (EV-AAV) vectors, was to harness their combinatorial potential of transient activity of cargo proteins and mRNA within EVs and relatively long-term expression of rAAV. The rigorous analysis of the EV-AAV particles produced with our method has shown minimal association between EV and rAAV particles, on average only 5% of total amount of rAAV genomes can be found in EV fractions. This result indicates that EVs primarily serve as a biological matrix during EV-AAV co-isolation and form a separate component in the final formulation. However, our data show that, despite the poor association with rAAV, the EV component can carry active molecules, such as mRNA or protein, which can be harnessed for delivering such short-lived activity molecules. Thus, our method based on EV and rAAV co-isolation utilizes similar properties of these two nanoparticle types. Our techniques allow for EV-AAV production and isolation, bypassing density gradient separation. This method involves continuous medium collection post-transfection, followed by polyethylene glycol (PEG)-mediated precipitation and differential centrifugation to obtain EV-AAVs. Incorporating fluorescently tagged CD9 into EV-AAVs allows for their direct detection post-isolation. Importantly, our findings demonstrate that EV-AAVs are full-fledged alternative to rAAV vector with the same biological function. EV-part in EV-AAV formulation does not show an adverse effect on mouse zygotes, in contrary EVs can carry short-lived molecules.

Overall, EV-AAVs represent a promising vector system with significant potential for both *in vitro* and *in vivo* applications, including transgenic mouse model generation and targeted gene delivery. They offer a versatile platform for exploring novel gene therapy strategies and addressing limitations associated with traditional AAV vectors. The following text introduces the developed method, presents preliminary data, and discusses the results gathered so far in the prepared manuscript titled "*Scalable Production of rAAV Vectors via Extracellular Vesicle-Mediated Purification for Gene Therapy and Transgenesis*".

Scalable Production of rAAV Vectors via Extracellular Vesicle-Mediated Purification for Gene Therapy and Transgenesis

Petr Nickl[1,2], Maria Barbiera[3], Jacopo Zini[3], Tereza Nickl[2], Aki Ushiki[4], Nadav Ahituv[4], Jan Prochazka[1,2], Marjo Yliperttula[3] and Radislav Sedlacek[1,2]

[1]*Czech Centre for Phenogenomics, Czech Centre of Phenogenomics - Transgenic and Archiving Module, Institute of Molecular Genetics of the Czech Academy of Sciences, Vestec, Czech Republic,*

[2]*Laboratory of Transgenic Models of Diseases, Institute of Molecular Genetics of the Czech Academy of Sciences, Vestec, Czech Republic,*

[3]*University of Helsinki, Division of Pharmaceutical Biosciences, Helsinki, Finland*

[4]*Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, California, USA*

[5]*Institute for Human Genetics, University of California San Francisco, San Francisco, California, USA.*

Abstract: This study introduces a novel method for the purification of adeno-associated viruses (AAVs) utilizing extracellular vesicles (EVs). By leveraging biologically active sub-populations of EVs, specifically exosomes and microsomes, as a matrix for the isolation and purification of AAV vectors, we developed a new formulation termed EV-AAV. This formulation can be used for both *in vivo* and *in vitro* delivery. We demonstrate that the EV-AAV vector efficiently delivers expression constructs and homology templates. Specifically, we assessed the effectiveness of EV-AAV vectors as carriers for homology templates using the CRISPR-READI method. Additionally, we utilized the EV-AAV vector to deliver the PiggyBac transposon system, evaluating its delivery efficiency and potential toxicity. Our results indicate that the delivery efficiency and toxicity profiles of the EV-AAV vector are comparable to those of standard AAV vectors. This work provides a detailed description of the production and characterization processes involved in generating EV-AAV vectors and explores their various downstream applications in transgenesis.

Highlights

- *Simplifies and streamlines the rAAV purification and delivery process while maintaining comparable efficacy to traditional methods.*
- *Enables high concentrations of rAAV in formulations, allowing for high-titer treatments with reduced adverse effects.*
- *Offers significant potential for both therapeutic applications and transgenic research, advancing gene delivery technology.*
- *Surpasses standard rAAV in delivery efficiency in vitro.*

Graphical abstract
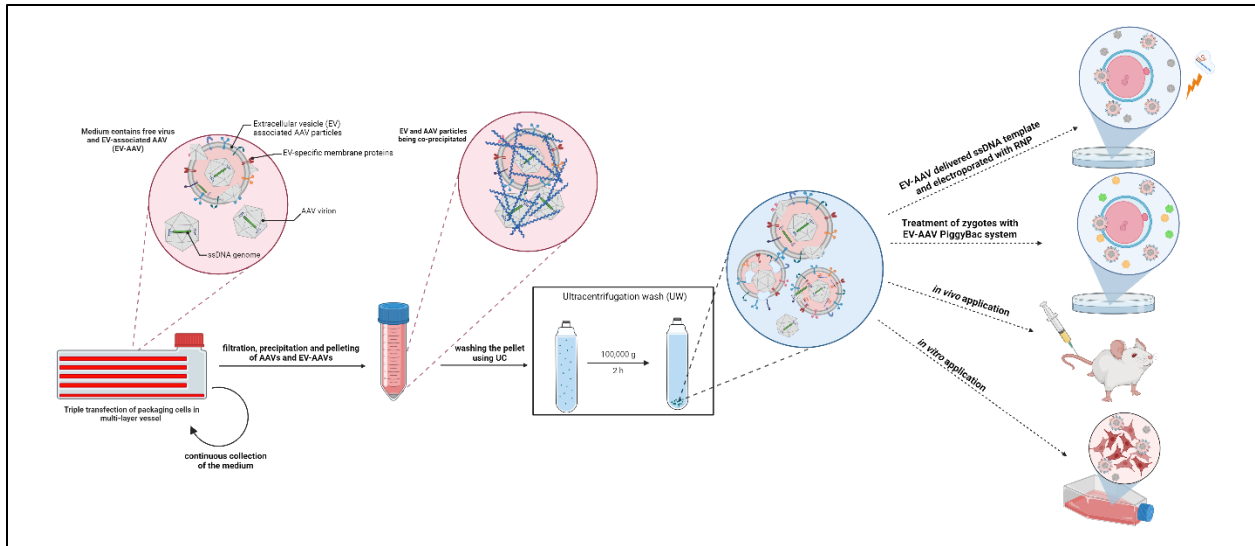
1. Introduction

Recombinant adeno-associated virus (rAAV) is a small, non-enveloped virus that has gained prominence in gene therapy due to its low immunogenicity and long-term expression capabilities. Despite its limited genome capacity of around 5kbp, rAAV is highly valued for its ability to achieve stable gene expression without integrating into the host genome. This characteristic reduces the risk of insertional mutagenesis, making rAAV a safer choice for therapeutic applications [6].

The production of rAAV in mammalian cells has evolved to bypass the need for a helper virus, which is necessary for wild-type AAV replication. This helper-free system involves the triple co-transfection of HEK293 cells with three essential plasmids, transfer, helper and Rep/Cap plasmids. Transfer plasmid contains the gene of interest (GOI) flanked by inverted terminal repeats (ITRs), which are crucial for packaging into the AAV virion. Helper plasmid replaces the need for a helper virus by providing adenoviral genes (E4, E2a, and VA) necessary for AAV replication. Rep/Cap plasmid encodes the replication proteins (Rep68, Rep52, Rep40) and capsid proteins (VP1, VP2, VP3) required for the AAV life cycle and packaging [174].

The helper-free method enables the production of high-titer rAAV vectors, which can be utilized for various applications. However, the isolation and purification of rAAV from the cell lysate pose significant challenges. Traditional methods such as affinity resins and ion exchange chromatography, while effective, are labor-intensive, time-consuming, and costly[172,173]. A commonly employed approach involves discontinuous gradient ultracentrifugation, which separates viral particles based on their density. Although effective, this method is complex, requiring specialized equipment and the removal of density gradient

medium [188]. The removal of the density medium is a crucial step in the isolation process, as it often leads to the aggregation of rAAV virions, resulting in non-functional virus and poor titer retrieval. Consequently, this makes the method less suitable for scalable concentration [189].

Extracellular vesicles (EVs) are proteolipid nanoparticles released by virtually all cell types into the extracellular environment. EVs form a heterogeneous population that includes *apoptotic bodies* (ABs), microvesicles, and exosomes, each varying in size and composition based on their origin and the pathways involved in their biosynthesis. *ABs* are produced by cells undergoing apoptosis. ABs range from 50 to 5000 nm and contain cellular organelles and degradative enzymes. They are formed through the blebbing of the plasma membrane. *Microvesicles* (50-1000 nm) bud directly from the plasma membrane and involve significant lipid and protein rearrangements. Unlike exosomes, their formation is not linked to the endosomal system. *Exosomes* are the smallest EVs (30-150 nm), exosomes originate from the endosomal system, specifically within multivesicular bodies (MVBs). Their biogenesis involves the inward budding of endosomal membranes and the formation of intraluminal vesicles (ILVs) [184]. Key proteins involved in exosome formation include tetraspanins (CD9, CD63) and ESCRT components (ALIX, TSG101). EVs play a crucial role in intercellular communication under both physiological and pathological conditions. Their ability to interact with cellular membranes and integrate into endogenous pathways makes them attractive as delivery vehicles for therapeutic agents [190,191].

This study introduces an innovative method for the purification of rAAVs using EVs, specifically exosomes and microvesicles, as a matrix for isolation and purification. This formulation, termed EV-AAV, simplifies the production process while maintaining the efficacy of traditional rAAV vectors. Our proposed method for EV-AAV production consists of three streamlined steps: *Continuous Medium Collection* - following the transfection of HEK293 cells with the three essential plasmids, the culture medium is continuously collected and replaced every 48 hours. This process, adapted from Benskey *et al.*, significantly increases the total viral yield without compromising virus quality [192]. *PEG-Mediated Precipitation* - the collected medium is filtered and undergoes precipitation using polyethylene glycol (PEG), a method that effectively concentrates both EVs and rAAVs from large volumes while preserving their physical and biological properties [193]. *Differential Centrifugation* - this step purifies the precipitated particles by removing insoluble protein/lipid aggregates, apoptotic bodies, and most of the residual PEG. The final formulation contains a mix of small EVs (50-300nm in size) and rAAVs. This method eliminates the need for density gradient ultracentrifugation, making the production process more straightforward and scalable. Additionally, to enhance EV production and facilitate particle characterization, we incorporated CD9 fused to mEmerald fluorescent protein into EV-AAVs, as shown by Schiller et al. 2018. This modification allows for the direct detection of exosomes/microvesicles after the final pelleting step.

EV-AAV vectors were tested for their ability to deliver expression constructs both *in vitro* and *in vivo*. The EV-AAV formulation of AAV serotype 9 (EV-AAV9) was tested on Neuro-2A and NIH3T3 cells to analyze the delivery potential of this compound formulation. Subsequently, the same formulation was tested *in vivo* to assess the stability and biodistribution of EV-AAV9 compared to rAAV9. The *in vitro* study revealed an improved delivery potential and short-term expression for the EV-AAV9 formulations. The *in vivo* bioactivity of the EV-AAV9 formulation demonstrated a similar short-term effect as EV-carried reporters in Lai et al. (2014)[195]. The results indicate that the EV-AAV formulation delivers active molecules (proteins, RNA) via the EV component and provides long-term expression through the rAAV component. This makes the EV-AAV9 formulation a versatile vector that combines transient activity with long-term stable expression.

In addition, EV-AAVs were also tested for their delivery potential to sensitive systems such as zygotes and embryos. Specifically, we evaluated the effectivity of EV-AAV1 (AAV serotype 1) vectors as carriers for homology templates, inspired by CRISPR-READI method [73]. Additionally, we utilized the EV-AAV vector to deliver the PiggyBac transposon expression system carrying MPRA (massively parallel reporter assay) reporters, assessing its delivery efficiency and potential toxicity in zygotes. Our results indicate that the delivery efficiency and toxicity profiles of the EV-AAV vector are comparable to those of standard rAAV vectors. Importantly, the presence of EVs does not adversely affect zygotes or embryos. On the contrary, EV-mediated purification allows for high concentrations of the virus in the formulation, enabling the use of high-titer treatments without the adverse effects typically associated with high rAAV titers.

In conclusion, the EV-AAV formulation offers a simplified and efficient method for rAAV purification and delivery, maintaining comparable efficacy to traditional rAAV vectors while streamlining the production process. This approach has significant potential for both therapeutic applications and transgenic research, representing a valuable advancement in gene delivery technology.

2. Materials and methods

Cell culture

293AAV cell Line (AAV-100) was purchased from Cell Biolabs, Inc., Neuro-2a cell line (CCL131) and NIHT3T cell line (CRL-1658) were purchased from American Type Culture Collection (ATCC). All cell lines were cultured in high-glucose Dulbecco's modified Eagle's medium-DMEM (Sigma-Aldrich) supplemented with 5% fetal bovine serum-FBS (Gibco, certified, United States) and penicillin (10,000U/mL)/streptomycin (10,000 µL/mL) solution (Gibco) in a humidified atmosphere with 5% CO_2 at 37°C.

Plasmids and constructs

AAV-CMV-Luc2 plasmid was generated by sub-cloning of synthetized luc2 (Photinus pyralis) CDS sequence into pAAV-GFP plasmid via EcoRI and HindIII restriction sites. AAV-CMV-mTom plasmid was generated by amplification of tdTomato reporter sequence from Rosa26 mT/mG (a gift from Liqun Luo, Addgene plasmid # 17787) with following primers F-5' AACATCGATTGAATTATTAAGGGTTCCGGATCCAT-3', R- 5' TGCTCGAGGCAAGCTCCCTCGAATCTCGAGTTACTTGTA-3'. The PCR product was inserted using In-Fusion® Snap Assembly Master Mix (Takara) into vector digested with pAAV-GFP (Cell Biolabs, AAV-400) via EcoRI and HindIII. HDR cassettes of pAAV-Lck loxP2_HDR, pAAV-Actn1 cKO_HDR, pAAV-Kcnma1-ctdTomato_HDR, pAAV-Tacstd2-IRES-CreERT2_HDR, pAAV-Nes-rtTA3-DTR-IRES-iRFP670-Far5, pAAV-Dpp4-rtTA3-DTR-H1-mKate2, pAAV-Ube3a-sl-BioID2_HDR, pAAV-Ube3a-ll-BioID2_HDR, were synthetized by GenScript, gene fragments were subcloned into pAAV-GFP backbone using NotI restriction sites. pscAAV-Sox2-mStr was a gift from Lin He (Addgene plasmid # 135617; http://n2t.net/addgene:135617; RRID: Addgene_135617). pAAV2/9n was a gift from James M. Wilson (Addgene plasmid # 112865). mEmerald-CD9-10 was a gift from Michael Davidson (Addgene plasmid # 54029; http://n2t.net/addgene:54029; RRID: Addgene_54029). pAAV-RC1 plasmid (VPK-421) and pHelper plasmid (part.no. 340202) were purchased from Cell Biolabs, Inc. PiggyBac system: hypertransposase and EGFP MPRA transposon plasmids were assembled and gifted by Nadav Ahituv and Aki Ushiki.

Standard rAAV production

Standard rAAV variant of the vector was produced by triple transfection in AAV293 cells seeded in one Corning® HYPERFlask® Cell Culture Vessels (Corning). PEI-mediated transfection (PEI: 24765-1, Polysciences, Inc.) was performed in Dulbecco's modified Eagle's medium (Sigma-Aldrich) supplemented with 2% fetal bovine serum (Gibco, certified, United States) and penicillin (10,000U/mL)/streptomycin (10,000 μL/mL) solution (Gibco). Per vessel total amount of plasmid DNA was 55μg in equimolar ratio 1:1:1 and incubated overnight. Used PEI:DNA ratio was 3:1. Next day, the medium was changed for DMEM (Sigma-Aldrich) supplemented with 2% FBS and penicillin/streptomycin solution (Gibco). After 4 days post transfection the cells and medium were collected and processed as described by [171]1999. Standard AAV vectors were stored at -80 °C (for long time storage) or 4°C when used the next day for other experiments. rAAV were resuspended in 1xPBS 0,001% Pluronic-F68 (24040032, Gibco).

EV-AAV production and isolation

Production of EV-AAV particles in Corning® HYPERFlask® Cell Culture Vessels (Corning) is initiated by triple transfection in AAV293 cells with AAV Rep/Cap, transfer and pHelper plasmids. PEI-mediated transfection was performed in Dulbecco's modified Eagle's medium (Sigma-Aldrich) supplemented with 2% fetal bovine serum (Gibco, certified, United States) and penicillin (10,000U/mL)/streptomycin (10,000 µL/mL) solution (Gibco). Per Hyperflask total amount of plasmid DNA was 50µg (for triple transfection) and 55µg (for transfection with additional pseudotyping plasmid, e.j. mEmerald-CD9-10) in equimolar ratio 1:1:1 (or 1:1:1:1) and incubated overnight. Used PEI:DNA ratio was 3:1. Next day, the medium was changed for DMEM (Sigma-Aldrich) supplemented with 2% EV-free FBS[2*] and penicillin/streptomycin solution (Gibco). First harvest was after 48 hours, since changing the medium to EV-free. Second and the third harvests were done in the same interval in respect to the previous medium change. The third and the last harvest includes collection of AAV293 cells from the vessel. The time scheme is graphically depicted in supplementary data (Supplementary method scheme). Immediately after collection, harvested medium was processed by filtering through PES Vacuum Filter 0.22 µm (431097, Corning) to remove cells and larger cellular debris. Filtered medium was mixed with 50% PEG6000 and 375 mM NaCl solution to obtain solution of 10% PEG6000 and final concentration of 75mM NaCl (NaCl in DMEM not considered). Solution was mixed gently my inverting a bottle and swirling. Once the solution properly mixed, samples were stored at 4°C up to 18 hours (overnight). Next day, medium was centrifuged in swing-bucket rotor at 3200G (maximum speed) and 4°C for 90 min. Supernatant was discarded and pellet was resuspended in 10ml of 1xDPBS and stored at -80°C. Once all harvests were collected and processed by precipitation, harvests after precipitation (3x10ml) were pooled and mixed with 45 ml of PBS. To remove cell debris, apoptotic bodies and larger microvesicles, the supernatant was centrifuged at 20,000xG and 4°C for 30min hour fixed-angle rotor (Eppendorf centrifuge). To pellet exosomes and AAV (particles >150nm), the supernatant from the previous centrifugation was subjected to 2h centrifugation at 100,000xG at 10°C using Ti70 in an Optima L-90K ultracentrifuge (Beckman Coulter, Indianapolis IN, USA). The EV-AAV pellet was resuspended in 0,2-0,5ml of serum-free 1xDPBS or PBS-HAT. Exo-AAV were stored at -80 °C (for long time storage) or 4°C and used the next day for other experiments. Production of rAAVs carrying HDR template can be performed with a single medium collection, average titer yield 1E+12 GC/mL is often sufficient for downstream treatment of zygotes *ex vivo* (Supplementary method scheme).

[2*] EV-free FBS was prepared by mixing FBS and DMEM medium in 1:1 ratio to make 50% FBS. This solution was ultracentrifuged overnight at 100,000xG and 4°C. Supernatant was collected and sterilized by filtration through 0.22µm syringe filter, then aliquoted and kept at -20°C.

Poly-ethylen glycol 6000 solution was prepared by dissolving Polyethylene Glycol 6000 (Fisher Chemical) and NaCl (S5886, Merck) powder in sterile-deionized water in w/v ratio. Solution was dissolved after constant stirring and elevated temperature 75-100°C for approx. 4 hours. The solution was sterilized by filtration through Vacuum Filter 0.22 µm (Corning) overnight and stored at 4°C. Final concentration of the PEG6000 NaCl solution waw 50% PEG6000 and 375 mM NaCl.

Western blot

Protein content of the samples was measured using BCA method by Pierce™ BCA Protein Assay Kit (Thermo Scientific, described below). 10 µg of each sample was prepared by mixing with 4x Laemmli Sample Buffer (BioRad) and boiling at 95°C for 5 min. The samples were separated with SDS-PAGE on a 4-20% Mini-PROTEAN® TGX™ Precast Protein Gel (BioRad) and transferred Trans-Blot Turbo Mini 0.2 µm Nitrocellulose membrane (BioRad) using Trans-Blot Turbo Transfer System (BioRad). Nitrocellulose membrane was stained with Ponceu S solution (BioReagent). The membrane was blocked with 10% (w/v) skimmed milk in Tris buffered saline-0.1% (v/ v) Tween-20 (TBS-T) for 30 minutes at room temperature and cut to strips corresponding to the weight of proteins to be analyzed. Each strip was then incubated with following primary antibodies diluted in 1% skimmed milk in TBS-T: anti-TSG101 (mouse), 1:1000 reducing conditions (BD Transduction Laboratories™ 612697); anti-HSP70 (mouse), 1:1000 reducing conditions (BD Transduction Laboratories™,610607); anti-CD9 (mouse), 1:1000 non-reducing conditions (HansaBioMed Life Sciences, Tallinn, Estonia, HBM-CD9-100); anti-α-tubulin (mouse), 1:1000, reducing conditions (Sigma Aldrich, DM1A); anti-CD9 (mouse), 1:2000 non-reducing conditions (Invitrogen,10626D); anti-GM130 (mouse), 1:500 reducing conditions (BD Transduction Laboratories™,610823), anti-AAV (rabbit), 1:1000 reducing conditions (Abcam 45482); rocking overnight at +4°C. The strips were then washed with TBS-T and stained with secondary antibody-HRP sheep Anti-mouse IgG (NXA931, 1:10,000, GE healthcare, Little Chalfont, UK) or goat anti-rabbit IgG (G-21234, 1:10,000, ThermoFischer Scientific) in TBS-T for one hour at room temperature. The strips were washed with TBS-T, assembled and incubated for 5 min with ECL substrate (Clarity™ Western ECL Substrate, BioRad) and imaged with ChemiDoc XRS (BioRad).

AAV quantification

AAV titer was quantified by real-time qPCR with LightCycler® 480 SYBR Green I Master Mix and primers that amplify AAV2 ITR sequence: F-5' GGAACCCCTAGTGATGGAGTT, R-5' CGGCCTCAGTGAGCGA by Aurnhammer *et al.* (2012) LightCycler 480 and related software (Roche) [196]. Before each quantification samples DNA was isolated and purified using High Pure Viral Nucleic Acid Kit (Roche).

Mass spectrometry

Mass spectrometry analysis was performed by a specialized external service facility (OMICS Mass Spectrometry Core Facility, BIOCEV). Due to the nature of this outsourcing, the specific details of the mass spectrometry protocol are not available for inclusion in this thesis. The facility used standard procedures and validated methods for the analysis, ensuring reliable and accurate results. All data obtained from the mass spectrometry have been appropriately analyzed. The amount 10 µg of protein of each tested (EV or EV-AAV) formulation was provided for mass spectrometry analysis in form of pellet obtained after 4 hours centrifugation at 20 000G, 10°C in 1.5 ml tubes

Nanoparticle Tracking Analysis (NTA)

Concentrations and size of particles were determined using Nanosight LM-14 instrument equipped with a 405 nm, 60 mW laser (Nanosight, Salisbury, Great Britain) and SCMOS cam- era (Hamamatsu Photonics K.K., Hamamatsu, Japan). The camera level was set to 15 and 60s of measurement time. The data were processed using NanoSight NTA software v3.0., with detection threshold adjusted to 5. Samples containing AAV were inactivated by heating at 75°C for 15 minutes prior to measurement.

Luciferase essay

Neuro-2A and NIH3T3 cells were seeded in 96-well Optical-Bottom plates (Thermo Scientofic, 160376), 20,000 cells per well in DMEM (Sigma-Aldrich) supplemented with 10% FBS (Gibco) and penicillin (10,000U/mL)/streptomycin (10,000 µL/mL) solution (Gibco). Next day, medium was changed for FBS free DMEM (Sigma-Aldrich) containing 50,000 MOI and 100,000 MOI of the std-AAV or EV-AAV. Estimated number of cells per well next day was 40,000, according to "Useful Numbers for Cell Culture" guideline by Thermo Scientific (https://www.thermofisher.com/cz/en/home/references/gibco-cell-culture-basics/cell-culture-protocols/cell-culture-useful-numbers.html). The cells were incubated with the viral samples overnight, then medium was changed for complete 10%FBS, DMEM, pen/strep and cultivated for another 48 hours until analyzed. Bioluminescence signal was measured using Varioskan™ LUX multimode microplate reader. Measurement of each well lasted 5s and each plate was read five times with 10 minutes between each measurement. Bioluminescence signal of each well was normalized to protein content of the particular well.

Bicinchoninic acid assay (BCA)

Protein content of all formulations (std-AAV, EV-AAV or EV-AAV-CD9) before SDS-PAGE was prepared and measured by Pierce™ BCA Protein Assay Kit (Thermo Scientific) protocol in Varioskan™ LUX multimode microplate reader.

Samples for protein content analysis after luciferase assay were prepared by removing growth medium from wells, washing each well with two times 100ul with warm 1xPhosphate Buffered Saline. Cell were lysed by adding 100μL RIPA Lysis and Extraction Buffer (Thermo Scientific) supplemented with Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific) and sonicated for 1 minute in Ultrasonic Bath (Bioblock). Lysates were further processed and measured by Pierce™ BCA Protein Assay Kit (Thermo Scientific) protocol in Varioskan™ LUX multimode microplate reader.

Animals and ethics statement

This study was performed in accordance with the Czech national laws and guidelines specifically by Animal Welfare Act No. 246/1992 Coll. The experiments were approved by the Committee on the Ethics of Animal Experiments of the Institute of Parasitology, Institute of Molecular Genetics of the Czech Academy of Sciences, and of the Departmental Expert Committee for the Approval of Projects of Experiments on animals of the Academy of Sciences of the Czech Republic (permit 101/2020). In this study, B6(Cg)-Tyr/J albino mice were used for *in vivo* experiments. Animals were always anaesthetized by intraperitoneal injection of ketamine and xylazine (0.1 mg per g body weight) (Biopharm) and 0.01 mg per g (Bioveta), respectively before intravenous injection. Zygotes were extracted from C57Bl/NCrl females.

Mice used for transgenic models production were bred in our specific pathogen-free facility (Institute of Molecular Genetics of the Czech Academy of Sciences; IMG). All experiments in this study were conducted in accordance with the ARRIVE guidelines and the laws of the Czech Republic. Animal protocol (93/2020) was approved by the Resort Professional Commission for Approval of Projects of Experiments on Animals of the Czech Academy of Sciences, Czech Republic.

All mice were euthanized using carbon dioxide (CO_2) inhalation followed by cervical dislocation. CO_2 inhalation was performed at a flow rate that displaced 30% of the chamber volume per minute to minimize distress.

Bioluminescence imaging of luciferase (Luc2) in vivo

Mice were injected intravenously through tail vein. First cohort of 20 mice, 4 mice per mixed group (6-8 weeks old) was injected with PBS, std-AAV $2,5x10^{11}$ GC/dose, EV-AAV $2,5x10^{11}$ GC/dose and EV-AAV-CD9 $2,5x10^{11}$ GC/dose. The cohort was measured for bioluminescent signal after intraperitoneal injection 300μL of XenoLight D-Luciferin - K+ Salt Bioluminescent Substrate (15 mg/ml; #122799, PerkinElmer) administration of each week for two months. Second cohort of 20 mice, 5 mice per mixed group (13-15 weeks old) was injected with PBS, std-AAV $2x10^{11}$ GC/dose, EV-AAV $2,5x10^{11}$ GC/dose. The cohort was measured for bioluminescent signal after intraperitoneal injection 300μL of XenoLight D-Luciferin - K+ Salt Bioluminescent Substrate each week for three months. Bioluminescence in each cohort was measured

with LagoX (Spectral Instruments Imaging) and analyzed with related software. During measurements the animals were anaesthetized by inhalation of isoflurane (Forane® Isoflurane 99.9%, Baxter).

RNA isolation and qRT-PCR

Mice in all cohorts were terminated by cervical dislocation and dissected. Their organs were further used for RNA isolation using TRI Reagent® (Sigma Aldrich). RNA was farther reverse-transcribed into cDNA and used for qRT-PCR. The reaction was performed with LightCycler® 480 SYBR Green I Master Mix and gene-specific primers amplifying Luc2 sequence, Luc2 F5'- GCTACAAACGCTCTCATCGACAAG, R5'- GTATTTGATCAGGCTCTTCAGCCG and reference gene Ubiquitin B (Ubb), F5'- ATGTGAAGGCCAAGATCCAG, R5'- TAATAGCCACCCCTCAGACG for normalization by [197],2001. Each organ sample was run in triplicate with each primer pair. RNA from the most affected organs was also analyzed without reverse transcription to exclude contamination by viral ssDNA genome in final analysis.

Multi Angle Light Scattering (MALS)

Eclipse NEON AF4 equiped with multiangle light scattering DAWN with 18 angles (Wyatt Technologies, Santa Barbara CA, US) and UV detector (Agilent Technologies, Santa Clara CA, US) used at 260 nm (not shown in the manuscript data). The analytical channel was equiped with regenerated cellulose membrane with 10 kDa MWCO, 400 um spacer (Wyatt, Santa Barbara CA, US).

Four different samples were tested: AAV, EV, EV-AAV (in triplicate) and mock (in triplicate). Mock samples consisted in 55 ul EVs (H1 luc2 ctrl) + AAV luc Pos 22 ul.

Table.1 Elution method

Mode	Duration (min)	Cross Flow start (ml/min)	Cross Flow end (ml/min)
Elution	1	0	0
Elution	1	1	1
Focus	2	1	1
Focus Inject	4	1	1
Focus	2	1	1
Elution	40	1	0.1
Elution	5	0.1	0.1
Elution	5	0.1	0
Elution	5	0	0
Elution	2	1	1

Raman spectrometry

The Timegate PicoRaman M3 (Timegate Instruments Oy, Oulu, Finland) with a 100 ps 532 nm laser and complementary metal oxide single-photon avalanche diode was connected to an Olympus microscope via microprobe. 2ul of samples where placed on calcium fluoride slide and air dried, each samples was

measured 30 times along the edges of the dry droplet with an exposure time of 6 seconds, 40x objective and ~25mW laser power, as used by Zini et al (2022).

Zygote preparation

Five-to-eight-week old female C57BL/6NCrl mice (Charles River), were superovulated by intraperitoneal application of 5 IU of pregnant mare serum gonadotropin (Calbiochem, Millipore, 367222), and 46–48 h later, 5 IU human chorion gonadotropin (Calbiochem, Millipore, 230734). Superovulated females were mated one-to-one with 5–25 weeks-old C57BL/6NCrl males to produce one-cell zygotes at 0.5 days post-coitum. Zygotes were collected and washed by Chen *et al.*, 2016. In short, zygotes were harvested from the ampulla of euthanized females, washed in hyaluronidase/M2 solution (Millipore, MR-051-F) to remove cumulus cells, washed four times in M2 media (Zenith, ZFM2–100) supplemented with 4 mg/ml bovine serum albumin (BSA, Sigma, A3311), and washed again four times in M2 + BSA media.

Zygote treatment by CRISPR-READI method

After zygote preparation, embryos were cultured in 20 mL droplets of KSOM + AA media (KCl-enriched simplex optimization medium with amino acid supplement, Zenith Biotech, ZEKS-050) containing corresponding EV-AAV vector dosage in 35 x 10 mm culture dishes (CellStar Greiner Bio-One, 627160) at 37°C with 95% humidity and 5% CO2 for 4 hours prior to RNP electroporation. EV-AAV-treated embryos were electroporated with assembled RNPs as previously described (Chen *et al.*, 2016). Briefly, embryos were transferred to 10 µL of Opti-MEM reduced serum media (Thermo Fisher, 31985062) with assembled Cas9/sgRNA RNPs. Embryos in RNP mixture were transferred to a 20mm length platinum plate electrode on glass slide, 1-mm gap, 1,5 mm height (Bulldog-Bio, CUY501P1-1.5) and electroporated (poring: 40V, pulse length – 3.5ms, interval – 50ms, number of pulses – 4, decay – 10%, + polarity; transfer: 5V wave, pulse length – 50ms, interval – 50ms, number of pulses – 5, decay – 40%, +/- polarity) using a (NEPA21, Nepagene). Zygotes were recovered from the cuvette by flushing three times with 100 mL of KSOM + AA media, then transferred into the culture droplets containing EV-AAV for a total incubation length of 24 hours. The following day, embryos were transferred to fresh KSOM + AA media overlaid with mineral oil until analysis or oviduct transfer. For generating live mice, treated zygotes that successfully developed into two-cell embryos were surgically transferred into the oviducts of pseudopregnant CD1 females (Charles River, Strain 022), using up to 18 embryos per oviduct.

Zygote treatment with PiggyBac system

After zygote preparation, embryos were cultured in 30 mL droplets of KSOM + AA media (KCl-enriched simplex optimization medium with amino acid supplement, Zenith Biotech, ZEKS-050) containing corresponding EV-AAV or AAV vectors dosage in 35 x 10 mm culture dishes (CellStar Greiner Bio-One,

627160) at 37°C with 95% humidity and 5% CO2 for 24 hours. Treatment with PiggyBac system consists of hypertransposase and EGFP-MPRA-transposon. Embryos were always treated with both construct in either EV-AAV or AAV form. The following day, embryos were transferred to fresh KSOM + AA media overlaid with mineral oil until analysis or oviduct transfer. For generating live mice, treated zygotes that successfully developed into two-cell embryos were surgically transferred into the oviducts of pseudopregnant CD1 females (Charles River, Strain 022), using up to 18 embryos per oviduct.

Cryogenic electron microscopy

CryoEM samples were prepared by plunge freezing using a Leica GP2. 4uL of the sample was applied to a glow-discharged cryoEM copper grid (Quantifoil R 1.2/1.3, 300 mesh) and incubated for 10s at room temperature and 80% relative humidity. The excess liquid was automatically blotted with the blot time of 3s and the grid was vitrified by plunging into the liquid ethane cooled to -180°C.

CryoEM data were acquired on a Jeol JEM 2100Plus transmission electron microscope (TEM) equipped with a LaB6 electron gun and TVIPS XF416 CMOS camera, operated at 200 kV. CryoEM images were acquired in a semi-automated mode using SerialEM software1 at a pixel size of 1.4 Å with a 3s exposure and a total electron dose of 46.9 electrons/Å.

Fluorescence imaging of embryos

Blastocysts were captured with Zeiss Axio Zoom V.16 microscope (2000 ms exposure), using a 560-585 nm excitation wavelength and a 600-690 nm emission wavelength, at a total magnification of 125x. E15.5 embryos were captured with an exposure time of 180 ms, using a 480 nm excitation wavelength and a 509 nm emission wavelength, at a total magnification of 11,2x. All images were processed in ZEN 3.0 (blue edition) software.

Genotyping

Born mice after AAV and EV-AAV zygote treatment were genotyped 4 weeks after birth. Blastocysts of treated embryos were lysed and genotyped 5 days after treatment using DreamTaq™ Green Buffer (10X, Thermo Scientific), DreamTaq DNA Polymerase (5 U/µL, Thermo Scientific) and Deoxynucleotide Set, 100 mM (Sigma Aldrich). Genotyping was performed with following primers: Sox2-mStr, Ube3a-BioID2, Lck loxP2 and Actn1 cKO.

3. Results

AAV vectors are co-isolated with exosomes and microvesicles

Here we present a protocol for the continuous collection and isolation of EV-AAV particles. Using examples of EV-AAV (Luc2) and EV-AAV-CD9 (Luc2) formulations, we demonstrate how individual

collections (C) differ mainly in terms of EV yield and AAV titer (Fig. 1a, b). The data show that EV production in both tested formulations peaks in EV count and AAV titer at C2, compared to C1 and C3, which show lower titers. The difference in EV yield is visible in Fig. 1a by the size of pellets of each formulation and corresponding collection (Fig.1a).

Additionally, overexpression of the CD9-mEmerald construct (CD9) labels exosomes and microvesicles, enabling direct visual characterization of the pellet (Fig. 1a). Although the expression level of the CD9-mE construct decreases over time, it is strongly active during C1 and C2, enhancing the production of small-sized, exosome-like EVs, as shown in the particle size table in Fig. 1b.

EV-AAV (Luc2) and EV-AAV-CD9 (Luc2) formulations were further tested for the presence of selected exosomal/microvesicular markers such as TSG101, HSP70, CD9, and the negative marker GM130. All markers were confirmed in the tested formulations except for GM130, a Golgi apparatus marker abundant only in cell lysate (Fig. 1c).

Distinct EV-AAV formulations of two different serotypes (rAAV2/1 and AAV2/9) have been generated. Some of these formulations were modified with CD9 to enhance the exosomal count. Table 1 summarizes all produced and used EV-AAV formulations in this work and their basic characteristics, including AAV titer, EV particle count, and the mode and mean size of nanoparticles in the formulations.

Table 1: Quantitative characterization of EV-AAV formulations

vector	serotype (genome form)	AAV titer [GC/mL]
EV-AAV-Luc2	AAV2/9n (ss)	6.2E+11 ± 1.2E+11
EV-AAV-Luc2-CD9	AAV2/9n (ss)	6.9E+11 ± 9.7E+10
EV-AAV-Ube3a-sl-BioID2	AAV2/1 (ss)	2.5E+12 ± 2.7E+11
EV-AAV-Ube3a-ll-BioID2	AAV2/1 (ss)	5.4E+11 ± 7.8E+10
Sox2-mStrawberry	AAV2/1 (sc)	4.3E+13 ± 1.6E+13
EV-AAV Lck loxP2	AAV2/1 (sc)	4.5E+12 ± 1.5E+12
EV-AAV Actn1 cKO	AAV2/1 (ss)	7.1E+13 ± 1.4E+13
Nes-rtTA3-DTR-IRES-iRFP670-Far5	AAV2/1 (ss)	1.2E+13 ± 6.9E11
Dpp4-rtTA3-DTR-H1-mKate2	AAV2/1 (ss)	1.3E+13 ± 4.5E+12
Tacstd2-IRES-CreERT2-pA	AAV2/1 (ss)	4.5E+12 ± 8.6E+11
cKcnma1-tdTomato	AAV2/1 (ss)	1.2E+13 ± 4.6E+12
EV-AAV Hypertransposase	AAV2/1 (ss)	5E+12 ± 1.4E+12
EV-AAV EGFP MPRA transposon	AAV2/1 (ss)	5.3E+12 ± 1.1E+12

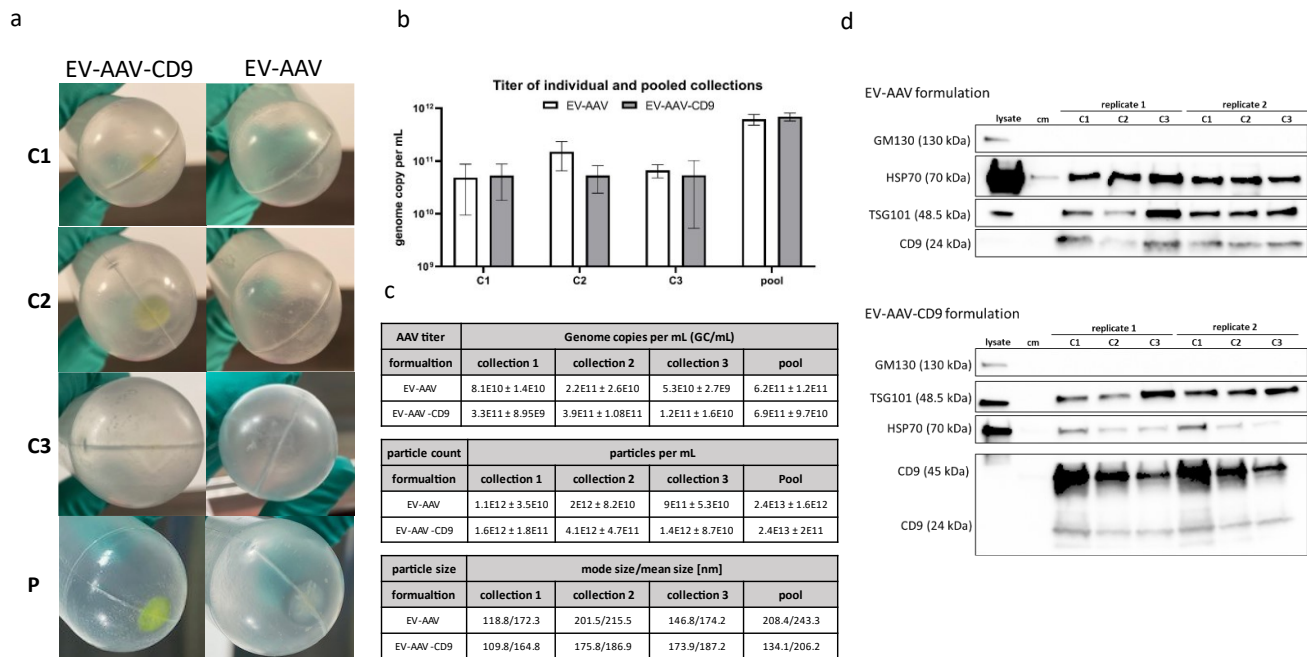ss – self-complementary AAV genome, sc – single-stranded AAV genome

AAV titer	Genome copies per mL (GC/mL)			
formualtion	collection 1	collection 2	collection 3	pool
EV-AAV	8.1E10 ± 1.4E10	2.2E11 ± 2.6E10	5.3E10 ± 2.7E9	6.2E11 ± 1.2E11
EV-AAV -CD9	3.3E11 ± 8.95E9	3.9E11 ± 1.08E11	1.2E11 ± 1.6E10	6.9E11 ± 9.7E10

particle count	particles per mL			
formualtion	collection 1	collection 2	collection 3	Pool
EV-AAV	1.1E12 ± 3.5E10	2E12 ± 8.2E10	9E11 ± 5.3E10	2.4E13 ± 1.6E12
EV-AAV -CD9	1.6E12 ± 1.8E11	4.1E12 ± 4.7E11	1.4E12 ± 8.7E10	2.4E13 ± 2E11

particle size	mode size/mean size [nm]			
formualtion	collection 1	collection 2	collection 3	pool
EV-AAV	118.8/172.3	201.5/215.5	146.8/174.2	208.4/243.3
EV-AAV -CD9	109.8/164.8	175.8/186.9	173.9/187.2	134.1/206.2

Figure 1: Characterization of EV-AAV formulations obtained by continuous collection
(a) EV-AAV pellets after ultracentrifugation wash for collection fractions (C1, C2, C3); P represents pooled and pelleted collections from C1, C2, and C3. (b) Absolute quantification of AAV viral titer across collections and pooled pellet. (c) Measured AAV titer, particle count, and particle size (mean/mode). (d) Western blot analysis showing specific positive and negative exosome/microvesicular markers for EV-AAV and EV-AAV-CD9 formulations.

Protein composition of EV-AAV formulation is dependent on rAAV serotype and cargo

The composition of EV-AAV formulations varies depending on the rAAV serotype used and the packaged construct of interest. Raman spectrometry (RS) revealed significant differences between standard rAAV particles and EV-AAV formulations. Standard rAAV particles form a distinct population, differentiating themselves from both EV-AAV and EV-AAV-CD9 formulations (Fig. 2a). Detailed analysis of EV-AAV formulations showed spectral similarities among EV, EV-CD9 (produced without virus), and EV-AAV9 formulations (Fig. 2b, c). In contrast, EV-AAV1 and EV-AAV1-CD9 formulations formed separate populations, indicating unique compositions compared to EV, EV-CD9, EV-AAV9, and EV-AAV9-CD9 formulations (Fig. 2c).

Comparative analysis using mass spectrometry demonstrated a higher amount of exosomal markers in the EV-AAV9 formulation compared to the reference EV formulation (Fig. 2d). Conversely, in the case of EV-AAV1, exosomal markers were more prominent in the reference EV formulation. These findings align with the RS analysis, indicating that EV-AAV9 is more similar to plain EVs.

Notably, all EV-AAV1 formulations mentioned so far carried homology-directed repair (HDR) templates for AAV-mediated transgenesis, thus transcriptionally inactive genomes. More detailed analysis of EV-

AAV1 formulations carrying distinct expression (EXP) cassettes confirmed lower similarity to no-virus EVs and revealed differences in protein composition based on the packaged expression cassette (Fig. 2e).

Figure 2: Comparative compositional analysis of EV-AAV vectors
(a) Principle component analysis (PCA) of measured spectra comparing extracellular vesicles (EV), extracellular vesicles associated AAV serotype 1 and serotype 9 (EV-AAV1, EV-AAV9) and adeno-associated virus (AAV). (b) Principle component analysis of measured spectra comparing extracellular vesicles (EV), extracellular vesicles pseudo-typed with CD9 (EV-CD9), extracellular vesicles associated AAV serotype 1 and serotype 9 (EV-AAV1, EV-AAV9), and extracellular vesicles associated AAV serotype 1 and serotype 9 pseudo-typed with CD9 (EV-AAV1-CD9, EV-AAV9-CD9). (c) Advanced principle component analysis of measured spectra comparing extracellular vesicles (EV), extracellular vesicles pseudo-typed with CD9 (EV-CD9), extracellular vesicles associated AAV serotype 1 and serotype 9 (EV-AAV1, EV-AAV9), and extracellular vesicles associated AAV serotype 1 and serotype 9 pseudo-typed with CD9 (EV-AAV1-CD9, EV-AAV9-CD9). (d) mass spectrometry difference analysis of EV-AAV serotype 9 carrying HDR template and EV (without AAV) - EV-AAV9 (HDR)/EV, comparison EV-AAV serotype 1 carrying HDR template and EV (without AAV) – EV-AAV1 (HDR)/EV. (e) mass spectrometry difference analysis of EV-AAV serotype 1 carrying expression cassette (hypertransposon) and EV (without AAV) - EV-AAV1 (EXP)/EV, comparison EV-AAV serotype 1 expression cassette (MPRA EGFP transposon) and EV (without AAV) – EV-AAV1 (EXP)/EV.

EV-AAV formulations contain a minor fraction of EV-associated virions, with the majority consisting of free AAV virions.

Structural analysis using cryo-electron microscopy (CryoEM) revealed the presence of membranous particles, indicating extracellular vesicles of distinct sizes. However, the structural analysis also confirmed a sparse association between EV and AAV particles, only a few AAV virions were encapsulated in EVs or attached to the EV protein corona, with most AAV virions detected outside the EVs (Fig. 3a).

The interaction strength between extracellular vesicles (EVs) and adeno-associated virus (AAV) particles was evaluated using Multi-Angle Light Scattering (MALS, Fig. 3c) and the separation of EV-AAV formulations in a continuous density gradient (CDG, Fig. 3b). EV-AAV formulations and mock controls (rAAV mixed with EVs) were separated over time by density and size. The results show a significant count of small particles (25-50 nm) eluted around 13-16 minutes (Fig.3c) and larger particles ranging from 50 nm to 150 nm eluted in later fractions (Fig. 3c, time 20-36 min). Quantification of AAV genome copy confirmed that the MALS fraction with a high number of small particles (25-50 nm) corresponds to free rAAVs (Fig.3d,e). In later fractions, the viral count decreased in mock samples; however, in the case of EV-AAV, the AAV titre was increased in those fractions (Fig.3d-e,), suggesting a slight association with EVs (Fig. 3d) compared to EV and AAV mixture.

This association was also confirmed by separation in the continuous density gradient (Fig.3b). The vast majority of AAV genomes were measured in the 40-50% iodixanol fraction (density 1,20-1,27 g/mL), corresponding to free rAAV virions (Fig. 3b). On average, only 0.9% of rAAVs were associated with EVs in the EV-AAV formulation. EV-AAV formulations consistently showed elevated levels of rAAV genomes in longer retention times (MALS) and low-density fractions, 1-1,144 g/mL (CDG), suggesting a small population of EV-associated AAVs (Fig. 3b-e).

a

b

c

d
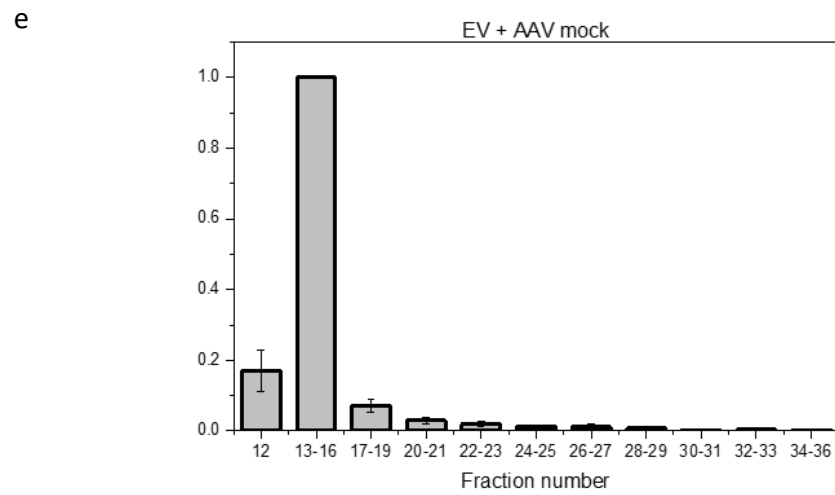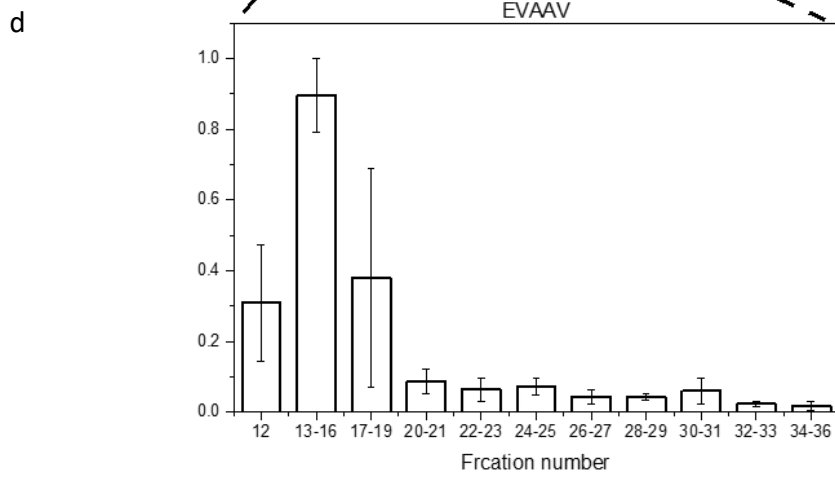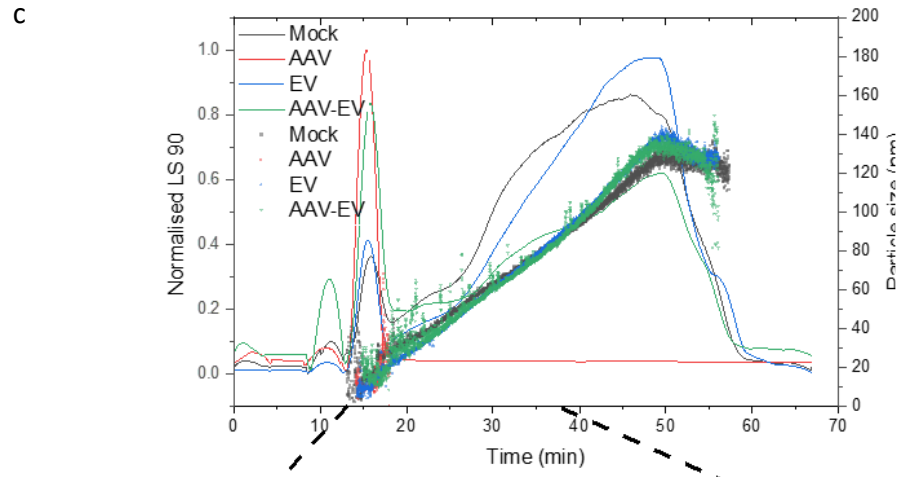

EVAAV

e

EV + AAV mock

146
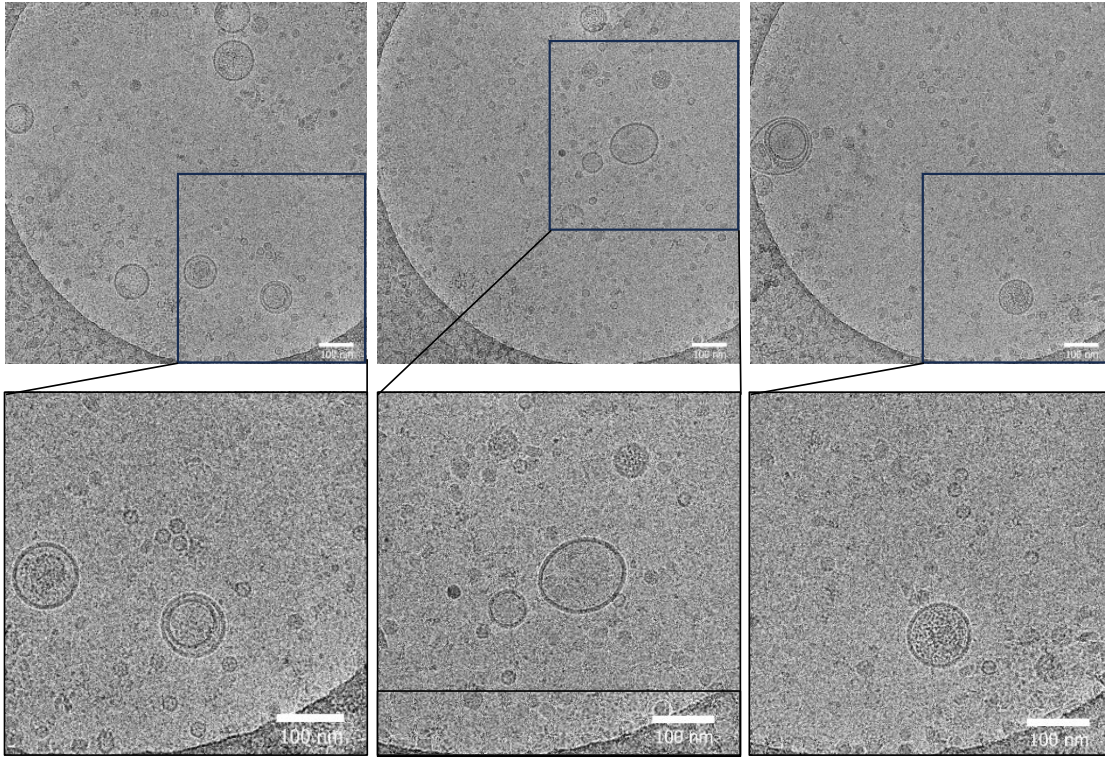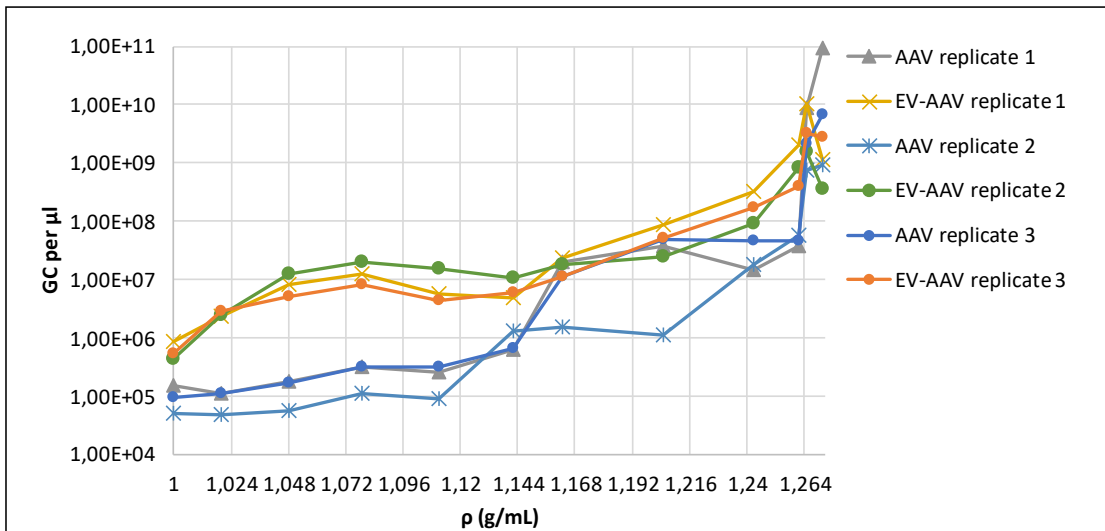
Figure 3: Capture of EV-AAV interactions and its detailed evaluation using Multi-Angle Light Scattering (MALS) and Continuous Density Gradient (CDG).
(a) CryoEM structural analysis confirms sparse EV-AAV association, with most AAVs located outside the EVs. (a) Quantification of AAV genomes in distinct fractions after CDG separation in an iodixanol gradient. CDG separates particles based on density: free rAAVs in 1.20-1.27 g/mL, while EVs are often found in 1.08-1.19 g/mL. (c) A4F/MALS elution profile of the light scattered at an angle of 90° (LS 90) and the particle diameter as function of elution time. LS90 axe is shown on the left and the data are represented as continuous lines, while the particle size is reported on the right axe and indicated as dots (black-Mock, red-AAV, blue- EV and green for EV-AAV). Distribution of AAV genome copies throughout fractions (d) distribution of AAV genomes in EV-AAV of A4F fractions +-1 SD of 3 runs (e) distribution of AAV genomes in mock of A4F fractions +-1 SD of 3 runs.

EV-AAV surpasses standard rAAV in delivery efficiency in treatment to Neuro-2A and NIH3T3 cells.

EV-AAV and rAAV vectors, both of serotype 2/9 (AAV9) and carrying the same luciferase (Luc2) gene, were compared for their delivery efficiency to the murine neuroblastoma cell line Neuro-2A (N2a) and the murine fibroblast cell line NIH3T3. The AAV9 serotype has a characteristic tropism for neural cells, making N2a cells, derived from the nervous system, more susceptible to AAV9 transduction. In the case of NIH3T3 cells, strong affinity for AAV9 has not been reported. In contrary, NIHT3T3 are considered less-permissive for rAAV transduction.

We evaluated the delivery potential of individual collections from EV-AAV and EV-AAV-CD9 formulations and standard rAAV, applying the same amount of virus across the tested formulations to both cell lines. In both N2a and NIH3T3 cells, the second EV-AAV collection outperformed other collections and standard rAAV, suggesting enhanced delivery efficiency due to the increased concentration of EVs in the second collection. The results were recapitulated while using lower MOI (Fig.4 a,b and c,d). The MOI positively correlates with intensity of the signal in EV-AAV treatment; however the trend is not shown in rAAV treated cells (Fig.4 a-d).

Comparative *In Vivo* Distribution and Stability of EV-AAV-Luc2 Particles and Standard AAV-Luc2 Vectors

To test the distribution and stability of EV-AAV-Luc2 (+CD9) particles compared to standard AAV-Luc2 vectors, the vectors were administered intravenously at a dose of 2.5E+11 CG per mouse (mice aged 6-8 weeks). The bioluminiscence signal was tracked weekly for 8 weeks in the first cohort, and then every 4 weeks in cohorts 2 and 3 (not shown). *In vivo* bioimaging data showed similar organ distribution of the bioluminiscence signal for all applied vectors, with strong initial expression in the liver. Over time, the signal weakened evenly in all groups and stabilized around week 4. Despite the same viral titers being administered, measurements a week after injection showed an increased signal in the rAAV groups across all cohorts (Fig. 5 and Fig. S1). This trend is illustrated in Fig.5, summarizing signal decline and stabilization between weeks 4 and 5 for each treatment group of cohort 1 (Fig.5b), the same trend was confirmed in percentual signal change in Fig. 5c.
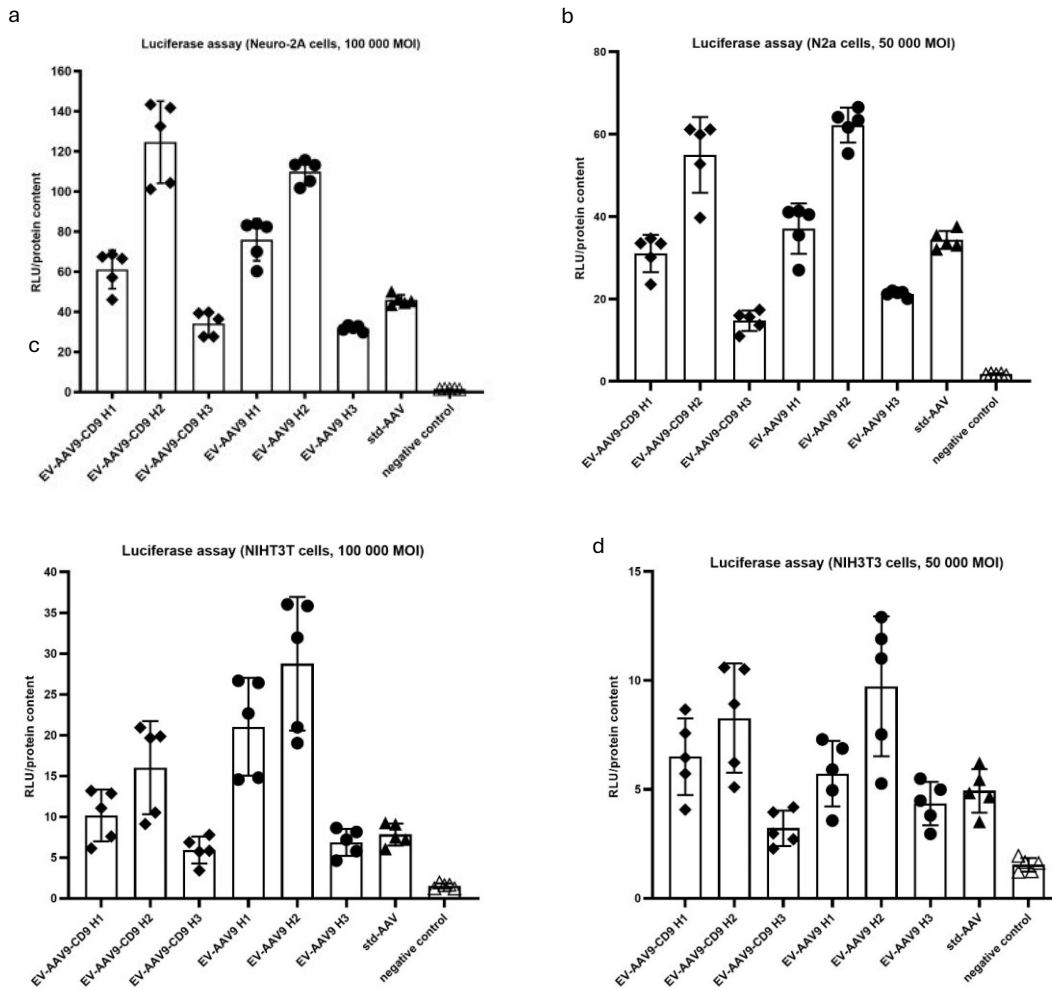
Figure 4: Evaluation of EV contribution to the transduction efficiency in two distinct cell lines (a) Bioluminiscence signal (RLU – relative luciferase unit) normalized to protein content in N2a cell treated with 100,000 MOI of EV-AAV and rAAV vectors. (b) Bioluminiscence signal (RLU) normalized to protein content in N2a cell treated with 50,000 MOI of EV-AAV and rAAV vectors. (c) Bioluminiscence signal (RLU) normalized to protein content in NIH3T3 cell treated with 100,000 MOI of EV-AAV and rAAV vectors. Bioluminiscence signal (RLU) normalized to protein content in NIH3T3 cell treated with 50,000 MOI of EV-AAV and rAAV vectors.

Final luciferase expression was evaluated at the end of the experiment by quantifying luciferase RNA levels across organs in cohorts 1 (Fig.5), 2, and selected animals in cohort 3 (Fig.S1). RNA analysis confirmed a lower level of luciferase RNA in the EV-AAV and EV-AAV-CD9 treated groups, with higher luc RNA levels in the EV-AAV-CD9 group compared to EV-AAV (Fig. S1g-i). Selected animals from cohort 3 were sacrificed at week 12, and their organs were imaged *ex vivo*, followed by Luc2 RNA level analysis. The results showed lower Luc2 signal and RNA levels in EV-AAV (CD9) treated animals copared to AAV group.

Overall, the analysis of all three cohorts demonstrated a similar biodistribution pattern dictated by AAV serotype 9, but lower reporter expression in the EV-AAV and EV-AAV-CD9 groups, despite the application of the same viral titer. These findings suggest that while the biodistribution is consistent, the stability and expression efficiency of EV-AAV formulations may be lower compared to standard AAV vectors.
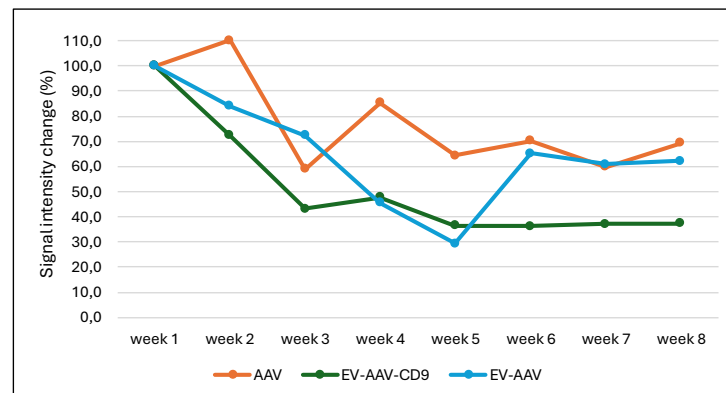

149
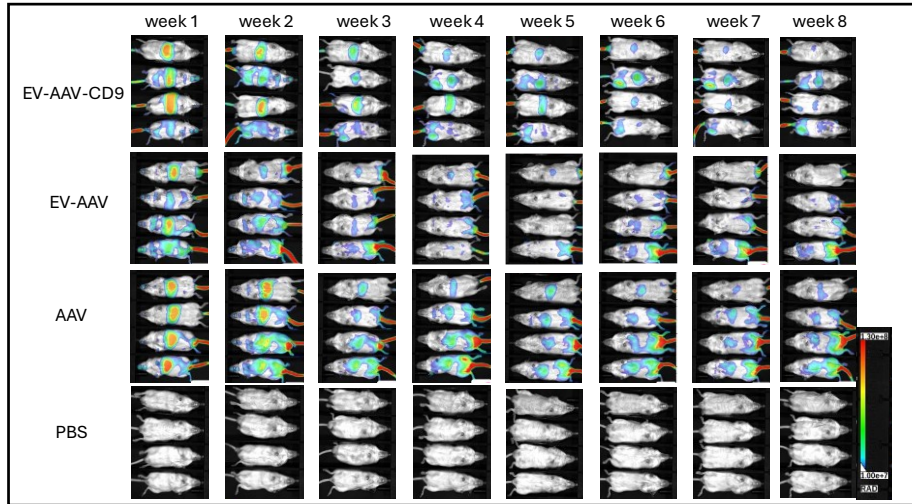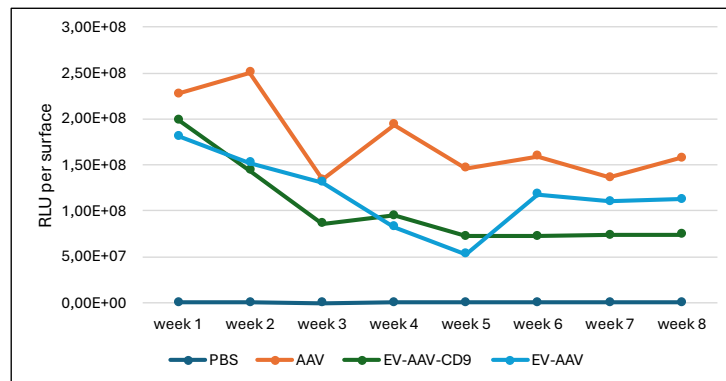
Figure 5: Comparative evaluation of bioluminescence signal and distribution *in vivo* (a) Bioluminiscence signal imaging of AAV, EV-AAV and EV-AAV-CD9 treated groups for 8 weeks with weekly acquisition. (b) Graph of intensity of bioluminescence signal in time. The signal is normalized to the surface of the animal. (c) Percentage of the signal change in time, graphical depiction of percentage change compared to the first measurement in week 1.

Versatile Application of EV-AAV Particles for CRISPR-Mediated Genome Engineering in Mice

The production of EV-AAV particles is not restricted by AAV serotype (as shown in Tab.1), allowing for the creation of vectors with varying tropisms. This versatility was demonstrated by generating EV-AAV1 vectors (serotype AAV2/1) carrying homology-directed repair (HDR) templates for site-specific genome engineering, facilitated by the CRISPR/Cas9 system. We constructed nine AAV transfer plasmid-based HDR templates within the AAV packaging limit. All nine constructs, packaged in EV-AAV, successfully delivered and integrated the transgene of interest.

Only animals with genomic insertions confirmed at both the 5' and 3' ends and complete transgene integration were considered positive. All nine transgenic strains confirmed site-specific insertion. Fig. 6a illustrates blastocysts treated with the Sox2-mStrawberry HDR vector, packaged in both AAV and EV-AAV. Both vectors successfully delivered the P2A-mStrawberry construct and inserted it into the Sox2 locus, resulting in the generation of functional Sox2-P2A-mStrawberry transgenic mice. Successful insertion is indicated by the red fluorescent signal in the inner cell mass, where Sox2 is expressed (Fig. 6a). The insertion in mice was further confirmed by specific amplification of the cassette at the genome-transgene boundary (Fig. 6b).

The nine transgenic strains were produced using seven single-stranded AAV vectors and two self-complementary AAV constructs. Optimal viral titers for zygote treatment were determined to maximize the number of transgenic animals while minimizing toxicity. Three different virus titers were tested: 5E+11 GC/mL (high), 1E+11 GC/mL (moderate), and 5E+10 GC/mL (low). High titer often leads to decreased number of born animals more lower number of positive animals to total born animals, and total transferred embryos (Fig.6c,d), however in several instances, high titers negatively impacted embryo viability, resulting in fewer or no births, suggesting toxicity at elevated titers. Conversely, low-titer treatments consistently produced positive animals, although at a lower percentage. Moderate EV-AAV titers yielded a consistent positive rate above 33% (Tab. 2).

Table 2 summarizes the applied titers, number of transferred embryos, and the number of born and positive animals, demonstrating the efficiency and toxicity of the applied doses. Data on genotyping and sequencing of the generated strains are provided in Fig. S6 and sequencing data in supplementary files (not disclosed in this work).

Table 2: Efficiency of targeted insertion using EV-AAV vectors

Transgenic construct delivered by EV-AAV	No. of transferred embryos			No. of born animals			No. of positive animals			Percentage of positive animals to born animals			Percentage of positive animals to transferred embryos		
Viral titer per µl (high/medium/ low)	high	med	low	high	med	low	low	med	low	high	med	low	high	med	low
EV-AAV-Ube3a-sl-BioID2 (ssAAV)	182	197	172	0	19	16	0	8	6	0	42,1	37,5	0	4,1	3,5
EV-AAV-Ube3a-ll-BioID2-CD9 (ssAAV)	136	-	-	13	-	-	5	-	-	38,4	-	-	3,7	x	x
Sox2-mStrawberry (scAAV)	64	64	75	3	0	13	1	0	6	33,3	0	46,2	1,6	0	8
EV-AAV Lck loxP2 (scAAV)	142	36	164	8	0	32	3	0	6	37,5	0	18,7	2,1	0	3,7
EV-AAV Actn1 cKO (ssAAV)	127	101	115	5	3	4	4	2	2	80	66,6	50	3,2	2	1,8
Nes-rtTA3-DTR-IRES-iRFP670-Far5 (ssAAV)	57	78	86	5	6	7	0	2	1	0	33,3	14,3	0	2,6	1,2
Dpp4-rtTA3-DTR-H1-mKate2 (ssAAV)	60	68	85	4	4	12	0	2	1	0	50	8,3	0	3	1,2
Tacstd2-IRES-CreERT2-pA (ssAAV)	19	34	34	0	5	11	0	2	6	0	40	54,5	0	5,9	17,6
cKcnma1-tdTomato (gene trap) (ssAAV)	58	56	53	4	6	6	2	3	1	50	50	16,6	3,4	5,4	1,9

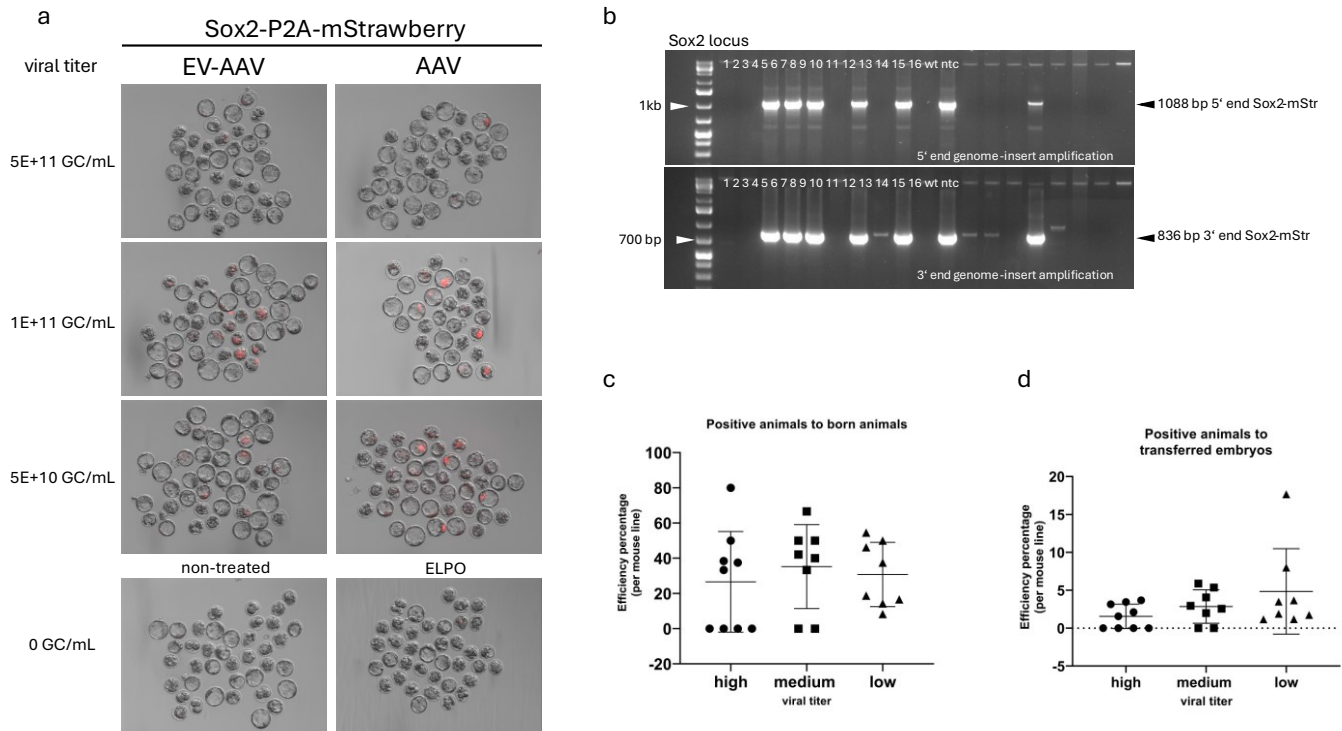high = 5E+11 GC/mL, medium = 1E+11 GC/mL, low = 5E+10GC/mL

Figure 6: Delivery of HDR template into zygote *ex vivo* using EV-AAV vectors
(a) Titration assay: Titration of EV-AAV and AAV vectors delivering the Sox2-mStrawberry transgene into the mouse genome using different vector titers. ELPO refers to samples electroporated only with Sox2 gRNA. (b) Agarose gel image: Specific amplification of the genomic-transgene boundary at both the 5' and 3' ends of the transgene, confirming site-specific insertion. (c) Graph of efficiency: A graph summarizing the efficiency of different EV-AAV titers, showing the percentage of positive animals relative to the total number of born animals. (d) Graph of embryo implantation efficiency: A graph summarizing the efficiency of different EV-AAV titers, showing the percentage of positive animals relative to the total number of implanted/transferred embryos into foster females.

Evaluating Toxicity, Delivery Efficiency and Replication Potential of EV-AAV Particles Using the PiggyBac System

The production of EV-AAV particles allows for concentration to high titers. However, high-titer treatments, and thus high multiplicity of infection, can have toxic effects on treated cells or tissues. To assess the maximum delivery potential and evaluate the toxicity of the expression and integrative system, the PiggyBac (PB) transposon system was employed to deliver the EGFP MPRA reporter into the mouse genome in as many copies as possible.

Additionally, we evaluated signal and copy number in the placenta, hypothesizing that some vectors might show a greater propensity for extra-embryonic tissue, allowing the embryo to sequester the vector in this protective tissue.

We compared the delivery and toxicity potential of rAAV and EV-AAV vectors. While both vector types were applied at the same viral titer, rAAV vectors achieved a higher integration rate. However, standard rAAV treatment resulted in a lower number of viable embryos (n=16) compared to EV-AAV treatment (n=36), suggesting increased toxicity associated with rAAV vectors.

In contrast, EV-AAV vectors, with a median integration rate of 2.62 copies per genome, yielded a higher number of viable embryos, thus increasing the median copy number per genome compared to rAAV (median = 2.45 copies per genome) (Tab. 3). These findings indicate that, although rAAV vectors can achieve higher integration rates, their associated toxicity reduces overall embryo viability. EV-AAV vectors, by comparison, offer a more viable option due to their lower toxicity and relatively high integration efficiency (Fig. 7a,b).

These results emphasize the need to balance integration efficiency with toxicity when selecting viral vectors for gene delivery. EV-AAV vectors present a promising alternative to standard rAAV vectors by providing a safer profile with fewer toxic effects, which is crucial for the successful development of transgenic models.

To further evaluate the toxicity limits of EV-AAV particles, a maximal dose of the vector was applied to mouse embryos. The doses used were 5E+12 GC/mL for the EGFP MPRA transposon and 1E+12 GC/mL for the hypertransposase. Due to limited rAAV titer availability, a comparison at these high doses was not feasible. The application of the maximal EV-AAV dose resulted in 60 viable embryos, with an average of 11.17 copies per genome and a median of 8.1 copies per genome (Tab. 3). These findings demonstrate that high-dose EV-AAV treatment can achieve substantial integration while maintaining embryo viability (Fig. 7c,d).

In summary, these findings suggest that EV-AAV vectors can be effectively used at high doses with fewer detrimental effects compared to rAAV vectors, making them a viable option for applications requiring high integration efficiency. The high titer and resulting high copy number per genome highlight the potential of EV-AAV vectors in genetic engineering. Fig.7e presents a comparison of signal intensity based on the copy number in the genome, along with the distribution of animals within groups of specific copies per genome.
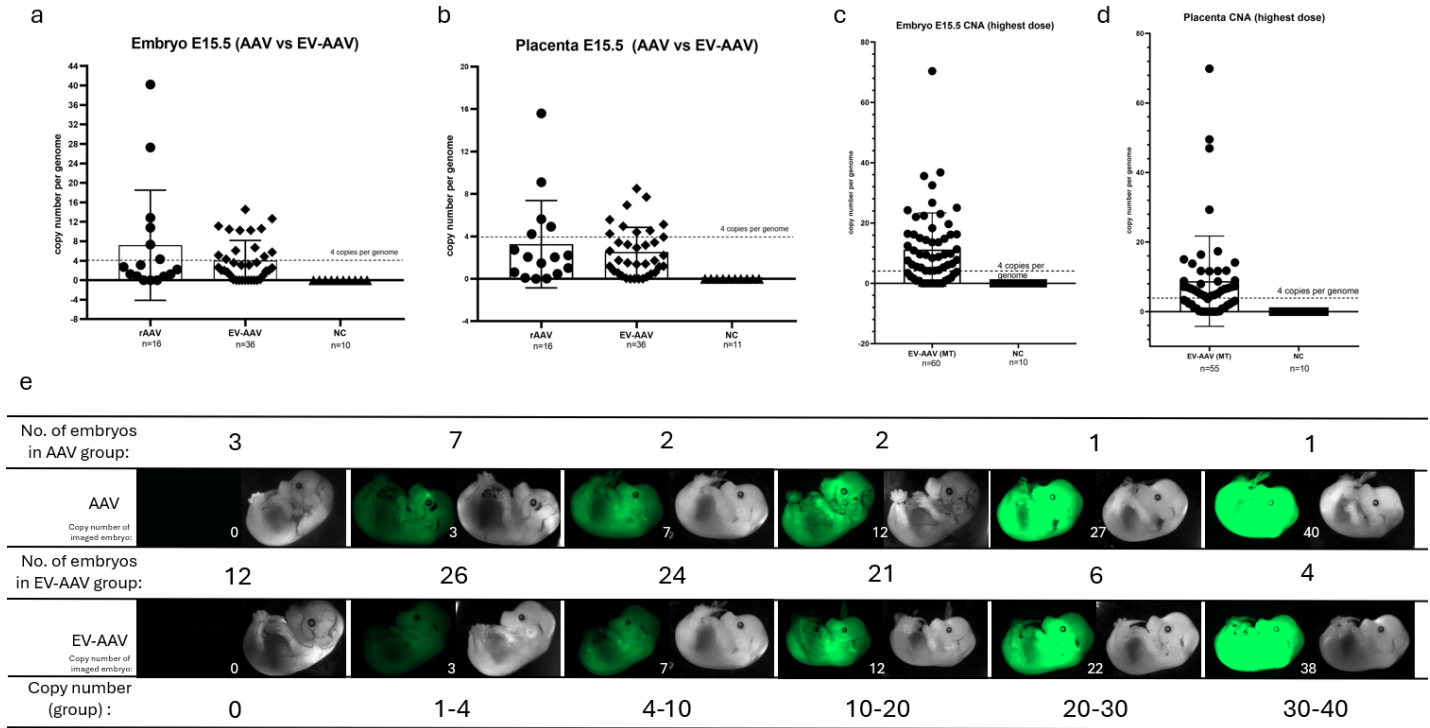
Figure 7: Delivery efficiency and toxicity evaluation of rAAV and EV-AAV PiggyBac vectors
(a) Graph comparing number of copies of PiggyBac system in the mouse genome delivered by rAAV and EV-AAV vectors in embryo, NC – non-treated (b) Graph comparing number of copies of PiggyBac system in the mouse genome delivered by rAAV and EV-AAV vectors in placenta (c) Graph showing number of copies of PiggyBac system in the mouse genome delivered by high dose of EV-AAV vector in embryo (d) Graph showing number of copies of PiggyBac system in the mouse genome delivered by high dose of EV-AAV vector in placenta. (e) Overview of rAAV and EV-AAV treated embryos (E15.5) and dependency of EGFP signal on number of PiggyBac copies in the genome. CNA – copy number analysis.

4. Discussion

Proper purification is essential for distinguishing particles of interest (viruses, EV-virus complexes, or EVs) from impurities in cell-conditioned media. Virus purification is well-established, typically using density gradient centrifugation, which separates particles based on sedimentation speed and density differences. Adeno-associated virus, composed mainly of proteins, have densities between 1.38 and 1.44 g/mL, higher than those of organelles and EVs.

Purifying EVs and EV-AAVs is more challenging due to the complexity of biological fluids or cell culture media. The chosen method depends on the biochemical composition, physical properties, volume, and study objectives. In the EV field, various methods are proposed. The MISEV2018 guidelines categorize methods based on their efficiency in separating EVs from non-EV material and yield [198]. The study and purification of virus-associated EV complexes, such as EV-AAVs, are relatively new. Density gradient centrifugation can distinguish between viral particles, free EVs, and associated EV-virus complexes. However, such separation requires a large number/concentration of input particles to obtain a relevant amount of all three separate subpopulations. Furthermore, each subpopulation is subjected to extensive analysis of particle count, size, and composition. In this study, we demonstrate that complete separation of viruses from extracellular vesicles might not be necessary, specifically for purposes of *ex vivo* gene delivery in mouse embryos.

We introduce a continuous EV-AAV production method using a transgenic AAV293 producer cell line. The procedure involves three collections over 48 hours. Each collection represents a separate EV-AAV batch suitable for downstream experiments. Individual collections differ in EV particle count and AAV titer but contain the same set of protein markers (Fig. 1b,c). These differences arise from culture development, with AAV production peaking between 72-96 hours post-transfection (collection 2 period, Fig. 1a). To increase the overall yield, collections were pooled, and further *in vivo* and *ex vivo* work was performed with these pooled formulations.

Analysis of formulation composition revealed the similarity of EV-AAV9 to plain EVs, whereas EV-AAV1 vectors show a unique composition, creating a separate subpopulation in principal component analysis derived from Raman spectrometry measurement (Fig. 2bc). This uniqueness in EV-AAV1 composition was also confirmed by mass spectrometry, showing a decreased number of exosome-specific proteins such as CD9, TSG101, CD81, and CD63 (Fig. 2d). Although MS analysis showed an elevated amount of adenovirus (Ad) single-stranded DNA binding protein (DBP) in the EV-AAV formulation, a protein associated with packaging and replication [50], no replication was detected in the treated embryos (Fig. S7, not disclosed in this work).

It has been reported that rAAV can associate with EVs during production in producer cells and remain associated during EV-AAV isolation using differential centrifugation protocols [199]. Our method aims for the isolation of EV-AAV using precipitation. Thorough analysis of the interaction between EVs and rAAVs after isolation with our method revealed a loose association of the two nanoparticles, averaging 0.9% of associated particles (Fig. 3). The formulations mainly consist of free EVs carrying proteins of interest (in the case of expression vectors, Fig. 2e) and free rAAVs carrying the gene of interest (GOI) in the form of an ssDNA genome.

154

The biological activity of EV-AAV complexes can be evaluated by comparing their activity with that of standard viruses and EVs. A luciferase assay comparing the short-term expression of EV-AAV and rAAV demonstrated significant delivery activity in the second collection of both EV-AAV and EV-AAV-CD9 formulations. The second collection of both EV-AAV formulations contained high concentrations of EV particles, which correlated with a strong luciferase signal in treated N2a and NIH3T3 cell lines, suggesting a short-term delivery capability associated with the EV component of the formulation (Fig. 4a-d). Further supporting this observation, different multiplicities of infection (MOI) in AAV-treated samples did not significantly elevate the luminescence signal (Fig. 4).

The *in vivo* bioluminescence assay revealed a decreased stability of the EV-AAV (CD9) signal after the first week, compared to rAAVs. Although all mice were injected with the same AAV titer (DNase I-resistant particles), all treated cohorts exhibited a reduced signal after one week. This decline is likely due to the lower functional stability of the EV-AAV formulation over time. Another possible explanation for this discrepancy is that EV-AAV (CD9) formulations contain DNase I-resistant particles in the form of either residual transfer plasmid encapsulated in EVs or rAAV. The stability of EVs carrying plasmids corresponds to the natural stability of EVs in mammalian systems, typically lasting a maximum of three days[200]. The *in vivo* bioluminescence assay tracking luciferase signal development showed a steady decline in luciferase activity in the EV-AAV-treated groups from the time of IV administration. In contrast, the AAV-treated mice demonstrated a less steep decline, characteristic of sustained AAV expression (Fig. 5)[201].

The EV-AAV1 formulation was used to deliver HDR templates into mouse zygotes, resulting in the generation of nine transgenic strains. The production of these strains revealed that the most efficient dose for achieving positive results was 1E+10 GC/mL, producing an average of 47% positive animals out of the total born. This dose, required for a 50 µL or 100 µL treatment, is easily achievable with the scalable concentration of EV-AAV formulations. The minimal required number of genome copies per 50 µL treatment drop is 2.5E+9 AAV GCs, which can be obtained in a single medium collection, allowing for a shortened production protocol (see Supplementary Methods).

EV-AAV1 proved to be a highly effective vector for delivering the PiggyBac (PB) transposon system in mouse embryos. EV-AAV was used to deliver transposons carrying the MPRA system to detect and localize non-coding regulatory elements (NCREs). The primary goal of the PB transposon system was to deliver as many copies of the MPRA reporter into the mouse genome as possible to study multiple NCREs in parallel, while minimizing detrimental effects on the embryo.

The EV-AAV formulation was compared to the rAAV variant of the PB system (Fig. 7ab). The titers for comparison were based on the maximal possible rAAV titer: 1E+12 GC/mL for the EGFP MPRA transposon and 3E+11 GC/mL for the hypertransposase. While rAAV-treated embryos and their placentas

showed higher copy numbers (Fig. 7ab), the number of viable and MPRA EGFP-positive embryos was higher in the EV-AAV group. Furthermore, only EV-AAV formulations allowed for the treatment of zygotes with significantly higher doses of PB vectors, thanks to the scalability of EV-AAV production. High-titer treatment resulted in an average of 11.17 copies per genome across 53 analyzed embryos (Fig. 7cd). The maximal dose treatment shows the delivery potential of EV-AAV of expression-active vector in high concentration into zygote or mouse embryo without significant toxic effect, which is common for standard rAAV. In addition, production of EV-AAV vectors allows scalable concentration of the rAAV virus, which is essential for treatments in such high concentrations.

5. Conclusion

This study demonstrates the potential of extracellular vesicle-associated adeno-associated virus (EV-AAV) as a versatile and scalable platform for gene delivery, particularly in the context of producing transgenic mouse models. Through a novel continuous EV-AAV production method, we achieved efficient isolation and functional delivery of EV-AAV formulations, offering a streamlined alternative to traditional rAAV production.

A key finding of this work is that complete separation of rAAV from extracellular vesicles (EVs) may not be necessary for effective gene delivery, especially in *ex vivo* applications, such as in mouse embryos. Our production method, which involves pooling multiple EV-AAV collections, yielded formulations with distinct particle and protein profiles that were biologically active. These EV-AAV formulations demonstrated effective gene transfer capabilities in both *in vitro* and *in vivo* experiments, despite differences in stability compared to traditional rAAV vectors. The results showed that the EV component of the formulations contributed to short-term delivery, while the rAAV provided sustained gene expression over time.

Furthermore, the scalability of EV-AAV production, combined with its ability to deliver high concentrations of genetic material, proved advantageous for generating transgenic strains. The EV-AAV1 formulation efficiently delivered transposon systems into mouse embryos, achieving high rates of transgenesis without significant toxicity, which is often a concern with standard rAAV treatments. The ability to concentrate EV-AAV formulations to high titers allowed for robust gene delivery, supporting the creation of multiple transgenic lines and enabling complex genetic studies.
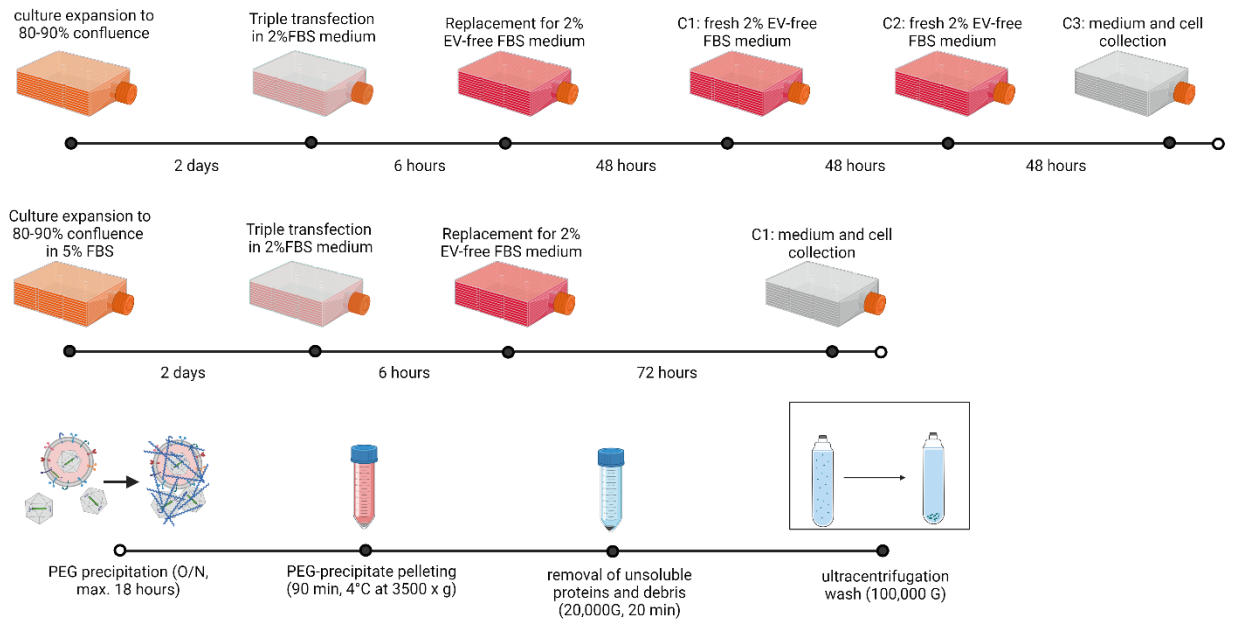
In summary, the EV-AAV system presents a promising alternative to traditional rAAV-based delivery methods. Its scalability, efficient gene transfer capabilities, and reduced toxicity make it a valuable tool for transgenic applications.

Acknowledgement

Supplementary data

Supplementary method scheme

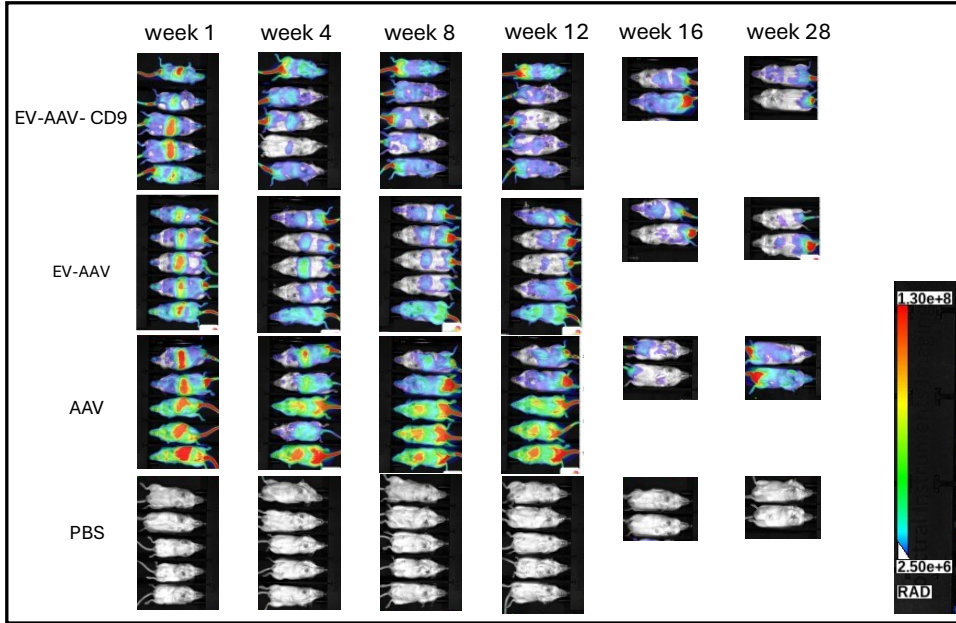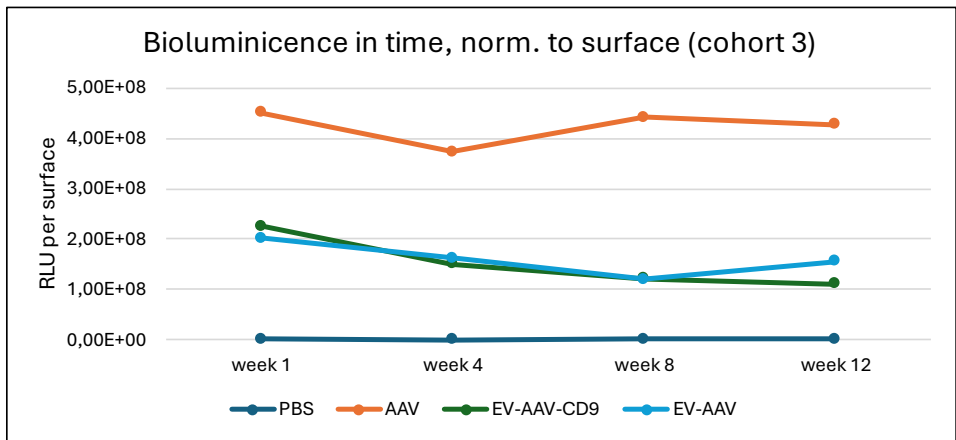The production of EV-AAV particles involves the expansion of AAV293 cells in a HyperFlask (multilayer vessel). Once the cells reach 80-90% confluence, the growth medium (5% FBS DMEM) is replaced with 2% FBS DMEM (transfection medium) by either fully replacing the growth medium or removing half of it and topping it off with 0% FBS DMEM to achieve 2% FBS. Cells are then transfected using the triple plasmid transfection protocol described in the methods section. Following 6 hours of incubation with the transfection mixture, the medium is replaced with 2% EV-free FBS medium (collection medium). The cells are cultured in this collection medium, and the medium can either be continuously harvested every 48 hours for three cycles or collected in a single harvest after 72 hours. Continuous harvesting typically results in higher EV-AAV yields. It should be noted that a single harvest often produces a sufficient titer for embryo treatment with the EV-AAV HDR vector. The collected medium is filtered through a 0.22 μm PES filter, and a 50% PEG6000 solution with 375 mM NaCl is added to obtain a final concentration of 10% PEG6000 and 75 mM NaCl. The medium-PEG-NaCl solution is thoroughly mixed and incubated overnight (18 hours). The next day, the medium is centrifuged at 3,500 G for 90 minutes, and the resulting pellets are resuspended in PBS. After resuspension, the solution can be stored at -80°C before further processing. The next step involves incubating the solution at room temperature for 10 minutes, followed by centrifugation at 20,000 G for 90 minutes at 10°C. The supernatant is then subjected to ultracentrifugation at 100,000 G for 2 hours at 10°C. The resulting pellet is resuspended in the buffer of interest. For long-term storage, resuspension in PBS-HAT is recommended, as reported in the literature by Gorgens et al. (2020) [202]. Created in Biorender.

157

a

2nd cohort, ventral view, 1min exposure time

b

Bioluminicence in time, norm. to surface (cohort 2)

c

Signal change in time (cohort 2)

158

d

| | week 1 | week 4 | week 8 | week 12 | week 16 | week 28 |

EV-AAV- CD9

EV-AAV

AAV

PBS

1.30e+8

2.50e+6
RAD

3rd cohort, ventral view, 1min exposure time

e

Bioluminicence in time, norm. to surface (cohort 3)

RLU per surface

5,00E+08
4,00E+08
3,00E+08
2,00E+08
1,00E+08
0,00E+00

week 1   week 4   week 8   week 12

PBS   AAV   EV-AAV-CD9   EV-AAV

f

Signal change in time (cohort 3)

Signal intensity change (%)

120
100
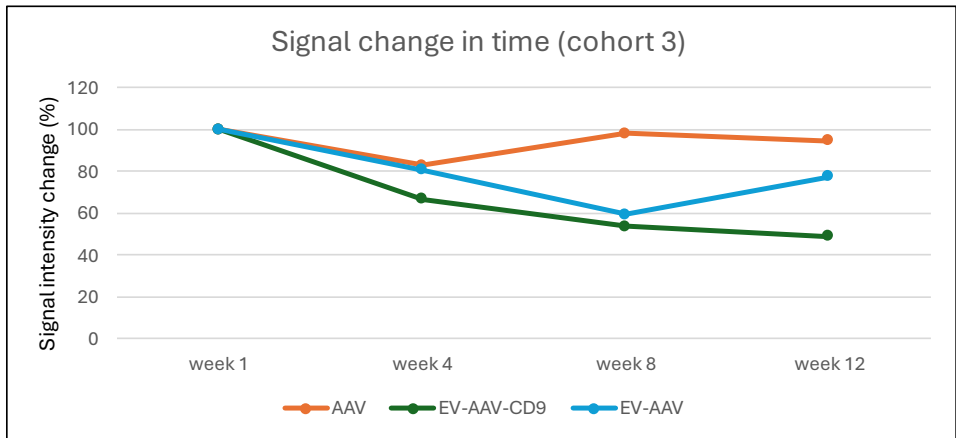80
60
40
20
0

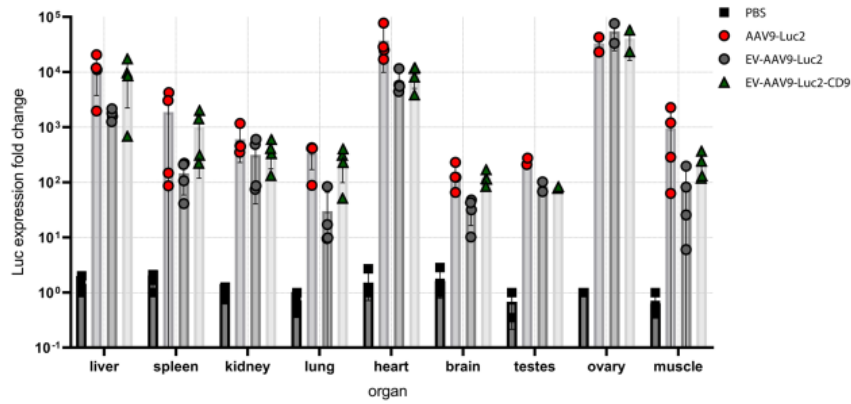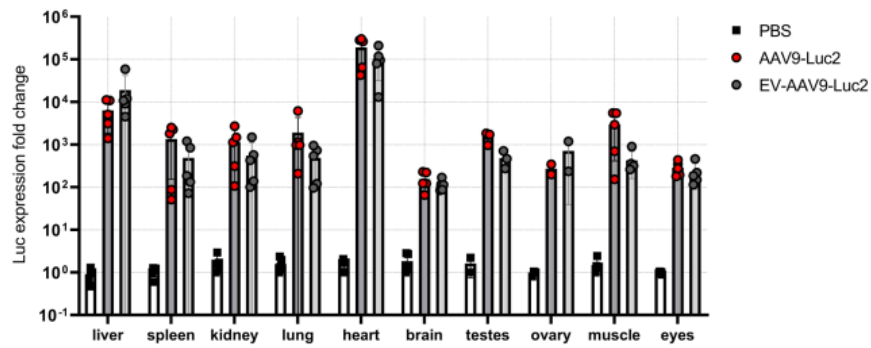week 1   week 4   week 8   week 12

AAV   EV-AAV-CD9   EV-AAV

g

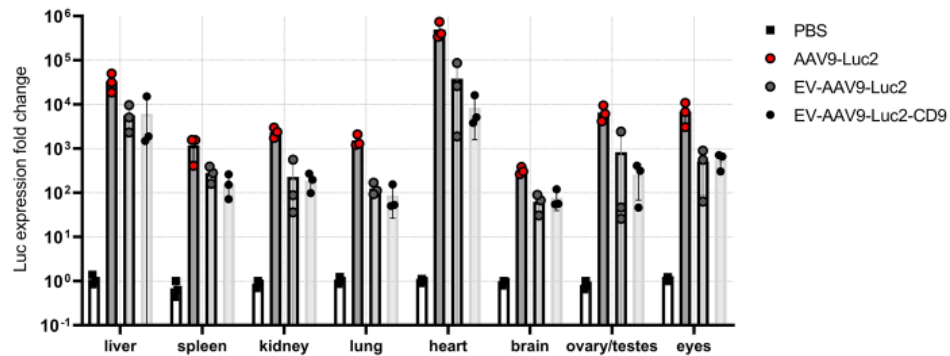luciferase expression in organs:cohort 1

h

luciferase expression in organs:cohort 2

i

luciferase expression in organs:cohort 3

Supplementary figure S1: Comparative evaluation of bioluminescence signal and distribution *in vivo*.
(a) Bioluminiscence signal imaging of AAV, EV-AAV treated groups (cohort 2) for 12 weeks with acquisition every 4 weeks. (b) Graph of intensity of bioluminescence signal in time. The signal is normalized to the surface of the animal. (c) Percentage of the signal change in time, graphical depiction of percentage change compared to the first measurement in week 1. (d) Bioluminiscence signal imaging of AAV, EV-AAV, EV-AAV-CD9 treated groups (cohort 3) for 12 weeks (selected animals for 28 weeks) with acquisition every 4 weeks. (e) Graph of intensity of bioluminescence signal in time. The signal is normalized to the surface of the animal. (f) Percentage of the signal change in time, graphical depiction of percentage change compared to the first measurement in week 1. (g) Relative expression of luc transgene in the analyzed organ by treatment group. Luc mRNA levels were normalized to internal reference gene Ubiquitin B gene (Ubb).

4.5 Summary

Aim 1

In Aim 1 of this thesis, we contributed to the development of a novel approach for generating a rapid disease mouse model. This model addressed the urgent need for accelerated research on SARS-CoV-2 and the progression of COVID-19 in organisms. It played a crucial role in testing bi-specific antibodies, which facilitated the neutralization of the SARS-CoV-2 virus. Additionally, we compared our transiently humanized model to other available senzitized and transgenic models in a review article, providing researchers with a resource to select the most appropriate model for their studies.

Aim 2

To facilitate and accelerate the production of mouse models based on site-specific recombinase systems, such as Cre/*loxP*, Dre/*rox*, FLP/*frt*, or Vika/*vox*, we developed and optimized an alternative method for direct allele conversion. This method, based on *ex vivo* rAAV treatment of zygotes, allows for highly efficient allele conversion without the need for extensive breeding and genotyping of the converted animals and their progeny. Additionally, this approach enables multistep sequential conversion in a single animal, significantly reducing the number of animals required and aligning with the 3R principles (Replacement, Reduction, and Refinement).

Aim 3

The production of rAAV is a laborious and tedious process. Unlike lentivirus, rAAV cannot be efficiently produced and purified through simple pelleting or precipitation. Our efforts have led to the development of a novel method based on the co-isolation of rAAV with extracellular vesicles to make this delivery platform more accesible. Specific AAV producer culture conditions enhance the production of extracellular vesicles, which serve as a biological matrix for rAAV collection. By harnessing the similar properties of these two distinct nanoparticles, we developed a straightforward and scalable method for rAAV production. The compound vector, EV-AAV (extracellular vesicle-AAV), has been tested in several applications, demonstrating either comparable or uniquely advantageous properties compared to traditional rAAV.

5. Discussion

This dissertation highlights the enormous potential of rAAV vectors for precise and efficient delivery, particularly in gene therapy. Despite their numerous advantages, rAAVs face several limitations, including packaging constraints, production inefficiencies, and integration risks. A deeper understanding of AAV genetics and its life cycle is crucial for improving rAAV vector production and application. This work provides a comprehensive overview of the molecular mechanisms involved in the interplay between the helper virus, host cell, and AAV.

Furthermore, the successful application of rAAV vectors depends on the correct selection of AAV serotype. Serotype plays a critical role in targeting specific tissues. For instance, transduction of zygotes or embryonic tissue is highly efficient with AAV2/1 or AAV2/6 serotypes but significantly less efficient with serotype AAV2/9 or others, as demonstrated by Chen *et al.* (2017)[73].

Another important factor in tissue-specific expression of rAAV vectors is the choice of promoter. Tissue- or cell-type-specific promoters enhance expression efficiency in the target tissue and reduce the risk of genotoxicity in off-target tissues[140,142]. Additionally, constitutive promoters, such as the cytomegalovirus promoter (CMV), can be subject to methylation-mediated silencing[203]. Therefore, strategic use of tissue-specific promoters is essential for targeted *in vivo* transduction.

The type of rAAV genome plays a crucial role in both titer yield during production and expression efficiency. A self-complementary genome (scAAV) carries mutated ITRs, allowing the packaging of both complementary strands within a single genome. This feature facilitates more efficient packaging due to its compact nature, with a packaging limit of 2.4 kb. Upon release in the nucleus, the double-stranded DNA (dsDNA) product is immediately ready for RNA transcription. In contrast, expression from a single-stranded AAV genome (ssAAV) requires second-strand synthesis or complementary genome annealing before transcription can occur. However, ssAAV genomes can carry larger transgene cassettes, up to 5 kb.

Thus, the choice between scAAV and ssAAV is driven by the size of the transgene and the need for rapid gene expression. Certain cell types, particularly non-dividing cells like neurons, transduce more efficiently with scAAV because they do not rely on cellular mechanisms to create a double-stranded template from single-stranded DNA.

Although rAAV vectors are generally considered non-integrative, their genome can integrate into the host genome at low frequency. The primary integration site for the AAV genome in humans is AAVS1 on chromosome 19, while in mice, the orthologous site is Mbs85 on chromosome 7[204]. AAVS1 and its orthologs in other species are associated with the persistence of the AAV genome in a latent state, which can be reactivated by helper virus infection. However, AAVS1 is not the only site prone to AAV genome

integration; other integration sites have also been identified. Off-target integration, particularly outside the AAVS1 locus, has been associated with the development of hepatocellular carcinoma[205]. In this context, it is crucial to analyze potential integration in studies where integration is not required, in order to avoid genotoxic effects.

5.1 Disease modelling *in vivo* using rAAV

The AAV-hACE2 vector used in the article *"Bispecific IgG Neutralizes SARS-CoV-2 Variants and Prevents Escape in Mice"* was designed as a single-stranded AAV (ssAAV) vector, carrying a CMV promoter, human β-globin intron, human ACE2 coding sequence (cDNA of the ENST00000252519.8 transcript derived from HEK293 cell mRNA), and a human growth hormone poly(A) signal (Extended Data Fig. 7a, ch. 4.1). The construct was packaged in an AAV2/9 serotype capsid. The rationale for this design was to achieve robust, universal expression of hACE2 in the target tissue, in this case, the lung. The AAV2/9 serotype efficiently transduces lung tissue (Fig.4, ch. 4.1), while the CMV promoter ensures strong, constitutive expression across all transduced cell types. The human β-globin intron enhances transcript processing, and the human growth hormone poly(A) signal ensures proper transcription termination.

The ACE2 receptor is recognized as the primary binding and entry factor for the SARS-CoV-2 virus in humans[206]. It has been shown that expressing this receptor alone can render cells that are not naturally susceptible to SARS-CoV-2 infection, such as murine cells, capable of viral entry[207]. Once SARS-CoV-2 crosses the cell membrane and enters the cytoplasm, it begins replication. While cell lysis can occur, the virus is also released through non-lytic mechanisms, leading to further infection[208].

The AAV-hACE2 vector was administered to the lungs using forced inhalation (FI), and analysis of the treated mice revealed high expression of the hACE2 protein in lung tissue (Extended Data Fig. 7b, ch. 4.1). This expression facilitated recognition of the cells by SARS-CoV-2 and subsequent viral internalization (Fig. 2a, ch. 4.1). In a separate, unpublished study using the same vector construct, but with the hACE2 coding sequence (CDS) replaced by the luciferase gene, we confirmed localized expression of the rAAV vector in the lungs (Fig. AAV-Luc vector expression in the lung after FI). This demonstrates that the route of administration can significantly contribute to the specific localization of transgene expression in the target tissue.

Although AAV-hACE2 vector enables rapid generation of a SARS-CoV2 sensitive mouse model. The expression of CMV-hACE2 construct is most likely occurring in cells that naturally do not express even mouse ortholog of ACE2. This is given by tropism of AAV2/9 serotype and constitutive character of CMV promoter activity. Therefore, in the context of this sensitized model it is important to emphasize that hACE2 expression pattern is driven by the serotype tropism and distribution of the virus in the tissue, and levels of

hACE2 are higher than endogenous expression of Ace2 due to robust CMV expression, as demonstrated in similar sensitized models[144,209,210].
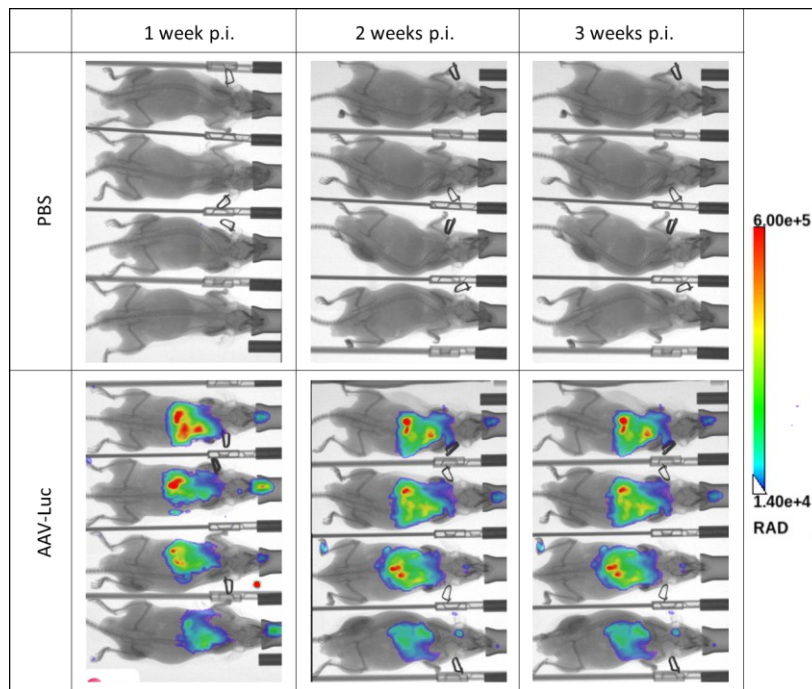
Figure: AAV-Luc vector expression in the lung after FI application
Development of bioluminescence signal throughout three weeks after forced inhalation, localization of the signal in lung and upper respiratory tract.

Despite the advantage of enabling rapid generation of an infection model on any mouse genetic background, sensitized models do not accurately reflect the natural expression patterns across cell types in the target tissue. Additionally, the expression of the delivered transgene is often higher than endogenous levels, depending on the promoter used. Therefore, their primary applications are in studying viral entry, replication[210], potential persistence, the inflammatory response[144,211], and testing specific reagents to prevent SARS-CoV-2 infection or severe COVID-19[143,209]. An additional unresolved question is the effect of the sensitizing vector itself on the inflammatory response.

In contrast to sensitized models, conventional transgenic models integrate the hACE2 gene into the genome under the control of specific promoters, enabling tissue- and cell-type-specific expression based on the promoter used and/or the genomic context (Fig.2, ch. 4.2). Transgenic models can be divided into two categories based on specificity and expression levels. Models with less tissue specificity and higher expression include the pCAGGS-hACE2 model, while more specific models, such as hACE2(loxP-STOP), rely on Cre drivers. The pCAGGS-hACE2 model drives strong, constitutive expression of hACE2 across various tissues[212]. In contrast, the hACE2(loxP-STOP) model is conditional, depending on the Cre driver, which can be tissue-specific or universal, allowing hACE2 expression in the desired tissue[213]. However,

these models face similar limitations to sensitized models in that expression is not confined to specific cell types, and the expression levels are relatively high.

More specific models, such as K18-hACE2[214], HFH4-hACE2[215], and Ace2-hACE2[216], mediate more endogenous and tissue-specific expression of hACE2, enabling a more precise study of the impact of SARS-CoV-2 infection on the organism, including systemic inflammatory responses and neuroinvasion. The most precise models in this context are those where the Ace2 promoter drives the expression of the hACE2 gene. The most advanced humanized model to date, developed by Zhang *et al.* (2023), replaces the entire mouse Ace2 gene with the human ACE2 ortholog, including its regulatory elements and introns. This model recapitulates the natural expression pattern of hACE2 in mice and allows for the study of long COVID-19 due to its moderate response to SARS-CoV-2, which is often observed in humans[217].

The formulation of the biological question and the focus of the study in the context of SARS-CoV-2 research should guide the choice of mouse model. Sensitized models are convenient for studying viral properties, including cell entry, replication, and virus neutralization. However, for systemic studies of SARS-CoV-2 infection—such as the systemic immune response, neuroinvasion, long COVID-19, or the infection's interaction with risk factors like age[218], diabetes, obesity[219], or hypertension[220,221]—transgenic models play a major role.

5.2   Application of rAAV in disease modelling

Our work, "*Multistep Allelic Conversion in Mouse Pre-Implantation Embryos by AAV Vectors*", describes the application of rAAV vectors for delivering site-specific recombinases (SSR) into mouse zygotes *ex vivo*. The primary goal of this study was to address challenges associated with SSR protein synthesis and provide the transgenic research community with a non-invasive method for efficient multistep allele conversion in mouse embryos. To achieve this, rAAV vectors were designed carrying a CMV promoter, human β-globin intron (HBB), codon-optimized site-specific recombinase coding sequence (CDS), and a human growth hormone poly(A) signal (Fig. rAAV SSR genomes). This construct was packaged in an AAV2/1 serotype capsid due to its relatively efficient production and its tropism toward embryonic tissues[136] (Fig. 4, ch. 4.3).

In this comprehensive study, we compared conventional methods of SSR delivery into zygotes, such as protein electroporation and mRNA pronuclear microinjection. Due to significant difficulties in synthesizing Flp and Vika proteins, we opted to validate Flp and Vika in the form of mRNA. Electroporation of SSR proteins with nuclear localization signals (NLS)[222] or pronuclear injection of mRNA[223] ensures either the direct delivery of active proteins or the direct translation of active proteins, driving transient SSR activity

in the early stages of embryo development (zygote or 2-cell embryo). This early-stage allele conversion results in efficient conversion across the majority of tissues with low mosaicism (Fig. 2b, ch. 4.3).

Delivery via rAAV vectors, however, is associated with a delay in gene expression within the host cell. As previously mentioned, the single-stranded AAV (ssAAV) genome requires complementary strand alignment or second-strand synthesis, resulting in delayed gene expression and SSR activity compared to the direct delivery of SSR proteins or mRNA[224]. To mitigate this limitation, we utilized only rAAV vectors expressing SSR coding sequences (CDS) under strong constitutive promoters (EF1a and CMV). This approach helps overcome the expression delay by ensuring the vector is delivered to all early embryonic cells, with the assurance of universal expression.
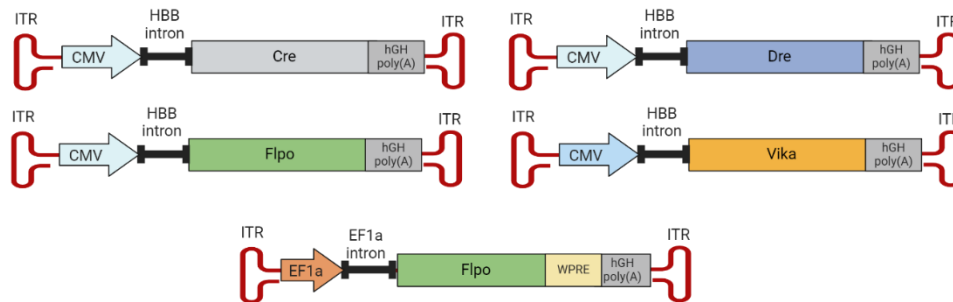

Figure AAV SSR genomes used in Multistep Allelic Conversion in Mouse Pre-Implantation Embryos by AAV Vectors publication. Abbreviations: CMV-cytomegalovirus promoter, HBB-human β-globin intron, hGH human growth hormone poly(A) signal. EF1a – human elongation factor 1 alpha, WPRE - Woodchuck Hepatitis Virus (WHV) Posttranscriptional Regulatory Element, ITR - inverted terminal repeat. Created in Biorender.

Titration of AAV SSR vectors in zygotes that developed into blastocysts revealed that certain AAV SSR vectors exhibit distinct efficiencies and impacts on viability at high titers, highlighting the effect of SSR expression on the overall success rate (Fig. 1, ch. 4.3).

The application of rAAV vectors for SSR delivery offers a significant advantage over protein- or mRNA-mediated conversion, particularly in the form of sequential multistep conversion. This process is based on the sequential incubation of zygotes with the corresponding viruses, allowing for two distinct conversion events, separated by the incubation period and the timing of AAV genome expression. For example, the combinatorial use of AAV-Flp and AAV-Cre recombinases can significantly accelerate the generation of a full knockout (tm1d) from the default tm1a state of the EUCOMM/KOMP knockout-first allele. This is achieved by progressing through the tm1c state (via AAV-Flp-mediated conversion) to the final tm1d state (via AAV-Cre-mediated conversion) (Fig. S26, ch. 4.3).

This study demonstrated the persistent nature of rAAV vectors from embryonic tissue to adult, as similarly reported by Deyle *et al.* (2009)[5]. Genotyping of the converted animals included an analysis of AAV genome persistence, either in episomal or integrated form, without distinguishing between the two. AAV genome persistence was confirmed in all AAV SSR vectors (Fig. S2c-S6c, ch. 4.3). Since the persistence rate was relatively low and some AAV genome-negative, converted mice were detected, we proceeded with further breeding of these animals. However, we report one instance of AAV genome integration, where the AAV genome was passed on to the next generation (Fig. S21, ch. 4.3). This highlights the necessity of genotyping for AAV persistence. In this study, AAV genome persistence was easily detected, allowing positive animals to be excluded from further experiments or breeding. However, detailed analysis of persistent AAV genome expression was not performed. Greig *et al.* (2023) reported that persisting genomes might suffer transgene expression loss.

The method of rapid, direct, and multistep allele conversion significantly reduces the time, cost, and number of animals needed to achieve specific genotypes. These advantages can promote the use of complex mouse models, such as EUCOMM/KOMP models or combinatorial reporter models, and deepen our understanding of multi-gene interactions. Moreover, similar to AAV-sensitized/humanized models, the application of AAV SSRs can potentially be used in mouse models with diverse genetic backgrounds. Both AAV-based platforms are accessible and affordable to the wider scientific community, helping to equalize research opportunities and and align with the 3R principles. This is made possible through the sharing of AAV transfer plasmids via non-profit plasmid repositories, which often provide rAAV production services.

5.3   rAAV Purification through the strategic utilization of EVs

The development of the isolation method based on extracellular vesicles (EVs) and the application of the resulting EV-associated AAV (EV-AAV) product originated from the goal of enhancing rAAV2/9 delivery efficiency across the blood-brain barrier (BBB)[199]. It has been reported that EVs or exosomes facilitate BBB crossing for rAAV vectors, potentially increasing overall transduction efficiency by shielding rAAV with an exosomal membrane[225]. Based on these findings, we aimed to produce similar compound vectors to deliver CRISPR-based editors targeting genes in the central nervous system (CNS).

During this process, we encountered several challenges. While many studies have reported relatively high yields of EV-AAV vectors produced in standard T150 or T175 culture formats, we were unable to replicate their efficiency[199,225,226]. To increase our yield, we scaled up the culture and established a production protocol using multilayer flasks (Hyperflask, Corning). Although we successfully increased the yield of EV-AAV vectors, isolating the formulation via conventional differential centrifugation from large volumes (550 mL) of medium proved labor-intensive.

Rider *et al.* (2016) reported an EV isolation method, ExtraPEG, based on PEG precipitation. Following their work, we optimized the precipitation protocol to concentrate rAAV, EVs, and EV-AAVs, while reducing the volume to work with. By combining the ExtraPEG method with the precipitation technique proposed by Guo *et al.* (2012), we established optimal precipitation conditions for both types of nanoparticles. This method and application of resulted compound vector is described in manuscript titled "*Scalable Production of rAAV Vectors via Extracellular Vesicle-Mediated Purification for Gene Therapy and Transgenesis*".

EV-AAV vectors produced using our method were subjected to particle analysis, a standard procedure for EV characterization. The EV component of the formulation was measured for size and particle number, and exosome-positive and exosome-negative markers were assessed using antibody staining. The AAV component was confirmed through antibody staining and titration. All data indicate the presence of both EV and AAV in the preparations. However, analysis of the interaction between the EV and AAV components revealed a weak association between the two. We also implemented CD9-GFP pseudotyping, as reported by Schiller *et al.* (2016), which yielded similar results, showing weak interaction between EV and AAV (Fig. 3a, b)[227].

Further *in vitro* testing of the compound vectors revealed that EV-AAV outperforms standard AAV in short-term delivery of a luciferase reporter construct (Fig. 4). Since rAAV expression is delayed due to the need for second-strand synthesis or complementary alignment, it takes longer for rAAV to express the carried transgene[228]. In contrast, EV-AAV vectors can deliver active molecules such as plasmids, proteins, and mRNAs, enabling them to produce the desired effect more quickly. The hypothesis that EVs can passively load proteins or mRNA from overexpressed genes in AAV producer cells was confirmed by anti-Cas9 staining (Fig. Protein analysis of EV-AAV formulations), further abundance of transgene product was also confirmed by mass spectrometry of EV-AAV- hypertransposase in Fig. Protein analysis of EV-AAV formulations and Fig.2e, ch. 4.4. The culture was producing an AAV vector carrying the gene encoding Cas9. However, it is important to note that the efficiency of passive loading is significantly limited compared to active loading methods, such as those mediated by lipid tags or proteins.[227]

In vivo testing of EV-AAV, pseudotyped EV-AAV (EV-AAV-CD9), and AAVs (all serotype 9), carrying a luciferase transgene did not replicate the superior results in EV-AAV-treated animals as reported by Li *et al.* (2023) and Gyorgy *et al.* (2014) (Fig. 5a, ch. 4.4). Our experiment did not include testing for neutralizing antibodies (NAb), as used in Li's report; however, even without NAb, our measurements did not align with their observations. The discrepancy could be attributed to the use of a crude pellet of EV-AAV in our experiment, which contained a mixture of free AAV, empty EVs, and a fraction of EV-AAV associations.

In contrast, Li *et al.* separated EV-AAV associations and tested only this strictly purified sub-population, which is more suitable for clinical translation[225,229].
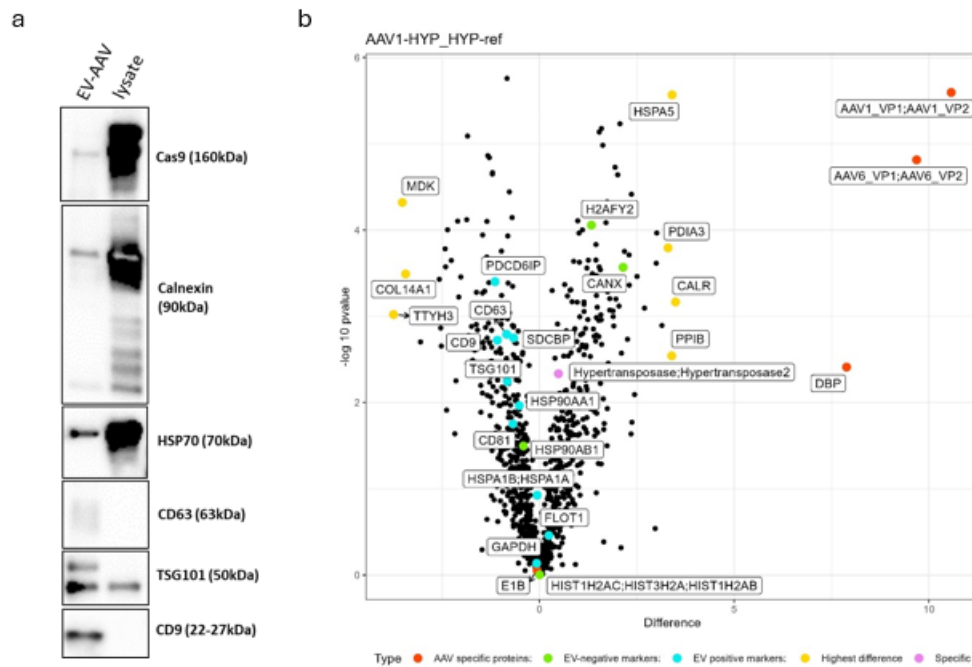

Figure: Protein analysis of EV-AAV formulations. (a)Antibody staining for exosomal markers in EV-AAV-Cas9 vector vs cells lysate:positive exosomal markers: HSP70, CD63, TSG101, CD9; negative marker: Calnexin; (b) and mass spectrometry analysis of protein composition of EV-AAV hypertransposase formulation: comparison of protein content between extracellular vesicles from cell overexpressing hypertransposase and EV-AAV vector from cells producing AAV and overexpressing hypertransposase.

To make a relevant comparison between EV-AAVs produced with our method and those produced by Li *et al.*, our EV-AAV preparations will be separated from AAV and other potential contaminants using a density gradient, as described by Li *et al.* (2023).

To evaluate the robustness and potential toxicity of the EV-AAV formulation, we tested the EV-AAV vector for treating mouse zygotes. Two types of vectors were used: one carrying a passive AAV genome and the other carrying a coding transgene, inspired by Chen *et al.* (2019) and Mizuno *et al.* (2018). The first type was used for the delivery of an HDR template for gene editing, while the second carried the PiggyBac system, consisting of two separate vectors—one expressing hypertransposase and the other carrying the MPRA EGFP transposon.

EV-AAV HDR vectors have proven to be a viable substitute for AAV vectors in treating sensitive systems such as mouse embryos. Despite using higher titers for the treatment of mouse zygotes compared to Chen *et al.* (2019) and Mizuno *et al.* (2018), we obtained similar HDR efficiency rates across all generated lines.

This establishes the EV-AAV method as a robust approach for producing HDR templates with high AAV titer yield and low toxicity.

Expression vectors were used to evaluate the toxic effects of the compound EV-AAV vector compared to AAV. The PiggyBac system was employed to integrate as many MPRA EGFP transposon copies as possible into the host genome. The presence of this genome-destabilizing factor can have a detrimental effect on treated embryos. Our data demonstrate, that AAV vectors are able to deliver more MPRA EGFP copies per genome at lower titers than EV-AAV. However, the number of animals with exceptionally high copy numbers is low in AAV treated group (Fig. 7, ch. 4.4). In contrast, EV-AAV vectors perform better at higher AAV titers. After treatment with a borderline toxic titer of 5E+12 GC/mL, EV-AAV-treated embryos exhibit higher viability, resulting in more high-copy animals that remain viable (Fig.7, ch. 4.4).

Since there is no existing literature on the application of EV or EV-AAV in the context of zygote treatment, we can only speculate on the contribution of the EV component in this biological system. It has been reported that EVs can cross the zona pellucida and play roles in processes such as gamete maturation, fertilization, and embryo implantation[230,231]. Therefore, we hypothesize that in our EV-AAV system, the EVs may reduce the toxic effects of AAV by potentially slowing down AAV internalization at the zygote membrane.

Based on our data, the EV-AAV formulation consists of separate nanoparticles with loose associations. Our initial hypothesis that AAV is either associated with or internalized within extracellular vesicles was incorrect. Future work will focus on utilizing the EV component as a transient vector carrying active molecules, such as proteins and mRNA, while the AAV component delivers passive HDR templates. Specifically, we aim to develop combinatorial particles where EVs are actively loaded with transient molecules (e.g., mRNA, proteins) and long-term AAV vectors provide HDR templates or auxiliary molecules to enhance the therapeutic effect. For example, EVs loaded with Cas9 and gRNA can induce a DNA double-strand break, while rAAV vectors supply the homology template for repair. This combined delivery approach holds great potential for targeted gene therapy.

6. Conclusion

This dissertation presents a series of innovative advancements in the application and production of recombinant adeno-associated virus (rAAV) for transgenic practice, addressing key challenges in model development, allele conversion, and scalable production. The overarching goal was to push the boundaries of rAAV's potential in genetic research, particularly in the creation of mouse models and the optimization of gene delivery systems.

A major contribution of this work was the development of a rapid, responsive mouse model for the study of SARS-CoV-2, offering a crucial tool during the global COVID-19 pandemic. This model allowed for efficient testing of bi-specific antibodies targeting the virus, thus enhancing the speed and accuracy of therapeutic evaluation. The novel system was also carefully compared with other transgenic models, providing a valuable resource to guide future research efforts on viral infections. This work highlighted the role of rAAV in addressing urgent, real-world challenges, emphasizing its flexibility in generating disease models tailored to emergent research needs.

Additionally, this thesis introduced a streamlined method for direct allele conversion in zygotes using *ex vivo* rAAV treatment. This approach significantly reduces the need for time-consuming breeding and genotyping, enabling efficient, site-specific genetic modifications in a single generation. The capacity for sequential allele conversion within the same organism also reduces the number of animals required, reflecting a commitment to the principles of Replacement, Reduction, and Refinement (3Rs) in animal research. This methodology offers a transformative tool for genetic research, facilitating faster and more precise model creation with fewer resources.

Another significant advancement was the development of a novel rAAV production method that leverages extracellular vesicles (EVs) as a biological matrix for co-isolation. This method provides a scalable and less labor-intensive alternative to traditional rAAV production techniques. The resultant EV-AAV compound vector demonstrated not only comparable but, in some instances, superior performance in various applications, illustrating the potential for this system to simplify production processes while maintaining efficacy. This work paves the way for broader applications of EV-based rAAV production, with potential implications for both research and therapeutic gene delivery.

Overall, this dissertation contributes to the growing body of knowledge on the use of rAAV in transgenic practices, offering innovative solutions to enhance efficiency, scalability, and ethical standards in genetic research. The methods developed in this work have the potential to significantly accelerate the production of genetically modified models and improve the scalability of gene delivery systems, ultimately benefiting a wide range of applications in both basic research and clinical settings.

7. References:

1. Atchison, R. W., Casto, B. C. & Hammon, W. M. D. Adenovirus-Associated Defective Virus Particles. *Science (1979)* **149**, 754–756 (1965).

2. Osten, P., Grinevich, V. & Cetin, A. Viral vectors: A wide range of choices and high levels of service. *Handb Exp Pharmacol* **178**, 177–202 (2007).

3. Kotin, R. M. *et al.* Site-specific integration by adeno-associated virus. *Proceedings of the National Academy of Sciences* **87**, 2211–5 (1990).

4. Kotin, R. M. & Berns, K. I. Organization of adeno-associated virus DNA in latently infected Detroit 6 cells. *Virology* **170**, 460–467 (1989).

5. Deyle, D. R. & Russell, D. W. Adeno-associated virus vector integration. *Curr Opin Mol Ther* **11**, 442 (2009).

6. Naso, M. F., Tomkowicz, B., Perry, W. L. & Strohl, W. R. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *Biodrugs* **31**, 317 (2017).

7. Ogston, P., Raj, K. & Beard, P. Productive replication of adeno-associated virus can occur in human papillomavirus type 16 (HPV-16) episome-containing keratinocytes and is augmented by the HPV-16 E2 protein. *J Virol* **74**, 3494–3504 (2000).

8. Atchison, R. W. The role of herpesviruses in adenovirus-associated virus replication in vitro. *Virology* **42**, 155–162 (1970).

9. Meier, A. F., Fraefel, C. & Seyffert, M. The Interplay between Adeno-Associated Virus and Its Helper Viruses. *Viruses* **12**, (2020).

10. Maurer, A. C. *et al.* The Assembly-Activating Protein Promotes Stability and Interactions between AAV's Viral Proteins to Nucleate Capsid Assembly. *Cell Rep* **23**, 1817–1830 (2018).

11. Collaco, R. F., Kalman-Maltese, V., Smith, A. D., Dignam, J. D. & Trempe, J. P. A Biochemical Characterization of the Adeno-associated Virus Rep40 Helicase. *Journal of Biological Chemistry* **278**, 34011–34017 (2003).

12. Santosh, V. *et al.* The Cryo-EM structure of AAV2 Rep68 in complex with ssDNA reveals a malleable AAA + machine that can switch between oligomeric states. *Nucleic Acids Res* **48**, 12983–12999 (2020).

13. Zarate-Perez, F. *et al.* Oligomeric Properties of Adeno-Associated Virus Rep68 Reflect Its Multifunctionality. *J Virol* **87**, 1232–1241 (2013).

14. Berthet, C., Raj, K., Saudan, P. & Beard, P. How adeno-associated virus Rep78 protein arrests cells completely in S phase. *Proceedings of the National Academy of Sciences* **102**, 13634–13639 (2005).

15. Di Pasquale, G. & Chiorini, J. A. PKA/PrKX activity is a modulator of AAV/adenovirus interaction. *EMBO J* **22**, 1716–1724 (2003).

16. Pereira, D. J., Mccarty, D. M., And, † & Muzyczka, N. The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. *J Virol* **71**, 1079–1088 (1997).

17. Beaton, A., Palumbo, P. & Berns, K. I. Expression from the adeno-associated virus p5 and p19 promoters is negatively regulated in trans by the rep protein. *J Virol* **63**, 4450–4454 (1989).

18. Schmidt, M., Afione, S. & Kotin, R. M. Adeno-Associated Virus Type 2 Rep78 Induces Apoptosis through Caspase Activation Independently of p53. *J Virol* **74**, 9441–9450 (2000).

19. Saudan, P., Vlach, J. & Beard, P. Inhibition of S-phase progression by adeno-associated virus Rep78 protein is mediated by hypophosphorylated pRb. *EMBO J* **19**, 4351–4361 (2000).

20. Batchu, R. B., Shammas, M. A., Wang, J. Y. & Munshi, N. C. Dual Level Inhibition of E2F-1 Activity by Adeno-associated Virus Rep78. *Journal of Biological Chemistry* **276**, 24315–24322 (2001).

21. Kyo¨stio¨, S. R. M., Kyo¨stio, K., Kyo¨stio¨, K., Wonderling, R. S. & Owens, R. A. Negative regulation of the adeno-associated virus (AAV) P5 promoter involves both the P5 rep binding site and the consensus ATP-binding motif of the AAV Rep68 protein. *J Virol* **69**, 6787–6796 (1995).

22. Han, S. I. *et al.* Rep68 protein of adeno-associated virus type 2 interacts with 14-3-3 proteins depending on phosphorylation at serine 535. *Virology* **320**, 144–155 (2004).

23. Smith, R. H. & Kotin, R. M. The Rep52 Gene Product of Adeno-Associated Virus Is a DNA Helicase with 3′-to-5′ Polarity. *J Virol* **72**, 4874–4881 (1998).

24. Agbandje-McKenna, M. & Kleinschmidt, J. AAV Capsid Structure and Cell Interactions. *Methods in Molecular Biology* **807**, 47–92 (2012).

25. Büning, H., Huber, A., Zhang, L., Meumann, N. & Hacker, U. Engineering the AAV capsid to optimize vector–host-interactions. *Curr Opin Pharmacol* **24**, 94–104 (2015).

26. Sonntag, F., Bleker, S., Leuchs, B., Fischer, R. & Kleinschmidt, J. A. Adeno-Associated Virus Type 2 Capsids with Externalized VP1/VP2 Trafficking Domains Are Generated prior to Passage through

the Cytoplasm and Are Maintained until Uncoating Occurs in the Nucleus. *J Virol* **80**, 11040–11054 (2006).

27.   Rabinowitz, J. E., Xiao, W. & Samulski, R. J. Insertional Mutagenesis of AAV2 Capsid and the Production of Recombinant Virus. *Virology* **265**, 274–285 (1999).

28.   Girod, A. *et al.* The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. *Journal of General Virology* **83**, 973–978 (2002).

29.   Chen, M. Y., Chen, W., Tong, J., Ho, M. L. & Suh, J. N-terminal serine/threonine motif has diverse and important effects on behavior of multiple AAV serotypes. *Virology* **563**, 107–115 (2021).

30.   Robinson, T. M. *et al.* An essential N-terminal serine-rich motif in the AAV VP1 and VP2 subunits that may play a role in viral transcription. *Virology* **546**, 127–132 (2020).

31.   Warrington, K. H. *et al.* Adeno-Associated Virus Type 2 VP2 Capsid Protein Is Nonessential and Can Tolerate Large Peptide Insertions at Its N Terminus. *J Virol* **78**, 6595–6609 (2004).

32.   Mietzsch, M. *et al.* Completion of the AAV Structural Atlas: Serotype Capsid Structures Reveals Clade-Specific Features. *Viruses 2021, Vol. 13, Page 101* **13**, 101 (2021).

33.   Büning, H. & Srivastava, A. Capsid Modifications for Targeting and Improving the Efficacy of AAV Vectors. *Mol Ther Methods Clin Dev* **12**, 248–265 (2019).

34.   Galibert, L. *et al.* Functional roles of the membrane-associated AAV protein MAAP. *Scientific Reports 2021 11:1* **11**, 1–19 (2021).

35.   Sonntag, F., Schmidt, K. & Kleinschmidt, J. A. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc Natl Acad Sci U S A* **107**, 10220–10225 (2010).

36.   Earley, L. F. *et al.* Adeno-associated Virus (AAV) Assembly-Activating Protein Is Not an Essential Requirement for Capsid Assembly of AAV Serotypes 4, 5, and 11. *J Virol* **91**, 1980–1996 (2017).

37.   Maurer, A. C. *et al.* The Assembly-Activating Protein Promotes Stability and Interactions between AAV's Viral Proteins to Nucleate Capsid Assembly. *Cell Rep* **23**, 1817–1830 (2018).

38.   Maurer, A. C., Cepeda Diaz, A. K. & Vandenberghe, L. H. Residues on Adeno-associated Virus Capsid Lumen Dictate Interactions and Compatibility with the Assembly-Activating Protein. *J Virol* **93**, 2013–2031 (2019).

39.   Cao, M., You, H. & Hermonat, P. L. The X Gene of Adeno-Associated Virus 2 (AAV2) Is Involved in Viral DNA Replication. *PLoS One* **9**, e104596 (2014).

40. Zolgensma | European Medicines Agency (EMA). https://www.ema.europa.eu/en/medicines/human/EPAR/zolgensma.

41. Matsushita, T. *et al. Adeno-Associated Virus Vectors Can Be Efficiently Produced without Helper Virus*. *Gene Therapy* vol. 5 http://www.stockton-press.co.uk/gt (1998).

42. Ryan, J. H., Zolotukhin, S. & Muzyczka, N. Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats. *J Virol* **70**, 1542–1553 (1996).

43. Snyder, R. O. *et al.* Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. *J Virol* **67**, 6096–6104 (1993).

44. Brister, J. R. & Muzyczka, N. Mechanism of Rep-mediated adeno-associated virus origin nicking. *J Virol* **74**, 7762–7771 (2000).

45. Richardson, W. D. & Westphal, H. Requirement for either early region 1a or early region 1b adenovirus gene products in the helper effect for adeno-associated virus. *J Virol* **51**, 404–410 (1984).

46. Carter, B. J., Antoni, B. A. & Klessig, D. F. Adenovirus containing a deletion of the early region 2A gene allows growth of adeno-associated virus with decreased efficiency. *Virology* **191**, 473–476 (1992).

47. Shi, Y., Seto, E., Chang, L. S. & Shenk, T. Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. *Cell* **67**, 377–388 (1991).

48. Chang, L. S., Shi, Y. & Shenk, T. Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. *J Virol* **63**, 3479–3488 (1989).

49. Chang, L. S. & Shenk, T. The adenovirus DNA-binding protein stimulates the rate of transcription directed by adenovirus and adeno-associated virus promoters. *J Virol* **64**, 2103–2109 (1990).

50. Ward, P., Dean, F. B., O'Donnell, M. E. & Berns, K. I. Role of the adenovirus DNA-binding protein in in vitro adeno-associated virus DNA replication. *J Virol* **72**, 420–427 (1998).

51. Piya, S. *et al.* The E1B19K oncoprotein complexes with Beclin 1 to regulate autophagy in adenovirus-infected cells. *PLoS One* **6**, (2011).

52. Matsushita, T. *et al.* The adenovirus E1A and E1B19K genes provide a helper function for transfection-based adeno-associated virus vector production. *J Gen Virol* **85**, 2209–2214 (2004).

53.   White, E. Regulation of the cell cycle and apoptosis by the oncogenes of adenovirus. *Oncogene* **20**, 7836–7846 (2001).

54.   Samulski, R. J. & Shenk, T. Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. *J Virol* **62**, 206–210 (1988).

55.   Stracker, T. H., Carson, C. T. & Weilzman, M. D. Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex. *Nature* **418**, 348–352 (2002).

56.   Allen, J. M., Halbert, C. L. & Miller, A. D. Improved adeno-associated virus vector production with transfection of a single helper adenovirus gene, E4orf6. *Mol Ther* **1**, 88–95 (2000).

57.   Nayak, R. & Pintel, D. J. Positive and negative effects of adenovirus type 5 helper functions on adeno-associated virus type 5 (AAV5) protein accumulation govern AAV5 virus production. *J Virol* **81**, 2205–2212 (2007).

58.   O'Malley, R. P., Mariano, T. M., Siekierka, J. & Mathews, M. B. A mechanism for the control of protein synthesis by adenovirus VA RNAI. *Cell* **44**, 391–400 (1986).

59.   Janik, J. E., Huston, M. M., Cho, K. & Rose', J. A. *Efficient Synthesis of Adeno-Associated Virus Structural Proteins Requires Both Adenovirus DNA Binding Protein and VA I RNA*. *VIROLOGY* vol. 168 (1989).

60.   Geoffroy, M.-C. & Salvetti, A. Helper functions required for wild type and recombinant adeno-associated virus growth. *Curr Gene Ther* **5**, 265–271 (2005).

61.   Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nature Protocols 2006 1:3* **1**, 1412–1428 (2006).

62.   Clément, N. & Grieger, J. C. Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Mol Ther Methods Clin Dev* **3**, 16002 (2016).

63.   Earley, L. F. *et al.* Adeno-Associated Virus Serotype-Specific Inverted Terminal Repeat Sequence Role in Vector Transgene Expression. *Hum Gene Ther* **31**, 151 (2020).

64.   Gao, G. *et al.* Clades of Adeno-Associated Viruses Are Widely Disseminated in Human Tissues. *J Virol* **78**, 6381–6388 (2004).

65.   Grimm, D. *et al.* Preclinical in vivo evaluation of pseudotyped adeno-associated virus vectors for liver gene therapy. *Blood* **102**, 2412–2419 (2003).

66.   Gao, G. P. *et al.* Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proceedings of the National Academy of Sciences* **99**, 11854–11859 (2002).

67. Mori, S., Wang, L., Takeuchi, T. & Kanda, T. Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsid protein. *Virology* **330**, 375–383 (2004).

68. Ogden, P. J., Kelsic, E. D., Sinai, S. & Church, G. M. Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science* **366**, 1139 (2019).

69. Lopez-Gordo, E., Chamberlain, K., Riyad, J. M., Kohlbrenner, E. & Weber, T. Natural Adeno-Associated Virus Serotypes and Engineered Adeno-Associated Virus Capsid Variants: Tropism Differences and Mechanistic Insights. *Viruses* **16**, 442 (2024).

70. Büning, H. & Srivastava, A. Capsid Modifications for Targeting and Improving the Efficacy of AAV Vectors. *Mol Ther Methods Clin Dev* **12**, 248–265 (2019).

71. Wu, Z., Asokan, A. & Samulski, R. J. Adeno-associated Virus Serotypes: Vector Toolkit for Human Gene Therapy. *Molecular Therapy* **14**, 316–327 (2006).

72. Issa, S. S., Shaimardanova, A. A., Solovyeva, V. V. & Rizvanov, A. A. Various AAV Serotypes and Their Applications in Gene Therapy: An Overview. *Cells 2023, Vol. 12, Page 785* **12**, 785 (2023).

73. Chen, S. *et al.* CRISPR-READI: Efficient Generation of Knockin Mice by CRISPR RNP Electroporation and AAV Donor Infection. *Cell Rep* **27**, 3780-3789.e4 (2019).

74. Qing, K. *et al.* Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. *Nat Med* **5**, 71–77 (1999).

75. Kashiwakura, Y. *et al.* Hepatocyte Growth Factor Receptor Is a Coreceptor for Adeno-Associated Virus Type 2 Infection. *J Virol* **79**, 609 (2005).

76. Akache, B. *et al.* The 37/67-kilodalton laminin receptor is a receptor for adeno-associated virus serotypes 8, 2, 3, and 9. *J Virol* **80**, 9831–9836 (2006).

77. Asokan, A., Hamra, J. B., Govindasamy, L., Agbandje-McKenna, M. & Samulski, R. J. Adeno-Associated Virus Type 2 Contains an Integrin α5β1 Binding Domain Essential for Viral Cell Entry. *J Virol* **80**, 8961 (2006).

78. Summerford, C., Bartlett, J. S. & Samulski, R. J. AlphaVbeta5 integrin: a co-receptor for adeno-associated virus type 2 infection. *Nat Med* **5**, 78–82 (1999).

79. Issa, S. S., Shaimardanova, A. A., Solovyeva, V. V. & Rizvanov, A. A. Various AAV Serotypes and Their Applications in Gene Therapy: An Overview. *Cells* **12**, (2023).

80. Zhang, H. *et al.* Several rAAV Vectors Efficiently Cross the Blood–brain Barrier and Transduce Neurons and Astrocytes in the Neonatal Mouse Central Nervous System. *Molecular Therapy* **19**, 1440 (2011).

81. Bell, C. L., Gurda, B. L., Van Vliet, K., Agbandje-McKenna, M. & Wilson, J. M. Identification of the galactose binding domain of the adeno-associated virus serotype 9 capsid. *J Virol* **86**, 7326–7333 (2012).

82. Pillay, S. *et al.* An essential receptor for adeno-associated virus infection. *Nature 2016 530:7588* **530**, 108–112 (2016).

83. Dudek, A. M. *et al.* GPR108 Is a Highly Conserved AAV Entry Factor. *Molecular Therapy* **28**, 367–381 (2020).

84. Pillay, S. *et al.* An essential receptor for adeno-associated virus infection. *Nature 2016 530:7588* **530**, 108–112 (2016).

85. Wu, Z., Miller, E., Agbandje-McKenna, M. & Samulski, R. J. Alpha2,3 and alpha2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *J Virol* **80**, 9093–9103 (2006).

86. Summerford, C. & Samulski, R. J. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J Virol* **72**, 1438–1445 (1998).

87. Ling, C. *et al.* Human Hepatocyte Growth Factor Receptor Is a Cellular Coreceptor for Adeno-Associated Virus Serotype 3. *Hum Gene Ther* **21**, 1741 (2010).

88. Di Pasquale, G. *et al.* Identification of PDGFR as a receptor for AAV-5 transduction. *Nat Med* **9**, 1306–1312 (2003).

89. Weller, M. L. *et al.* Epidermal growth factor receptor is a co-receptor for adeno-associated virus serotype 6. *Nat Med* **16**, 662 (2010).

90. Dhungel, B. P., Bailey, C. G. & Rasko, J. E. J. Journey to the Center of the Cell: Tracing the Path of AAV Transduction. *Trends Mol Med* **27**, 172–184 (2021).

91. Seisenberger, G. *et al.* Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science* **294**, 1929–1932 (2001).

92. Bartlett, J. S., Wilcher, R. & Samulski, R. J. Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J Virol* **74**, 2777–2785 (2000).

93. Sanlioglu, S. *et al.* Endocytosis and nuclear trafficking of adeno-associated virus type 2 are controlled by rac1 and phosphatidylinositol-3 kinase activation. *J Virol* **74**, 9184–9196 (2000).

94. Nonnenmacher, M. & Weber, T. Adeno-associated virus 2 infection requires endocytosis through the CLIC/GEEC pathway. *Cell Host Microbe* **10**, 563–576 (2011).

95. Weinberg, M. S. *et al.* Recombinant Adeno-Associated Virus Utilizes Cell-Specific Infectious Entry Mechanisms. *J Virol* **88**, 12472 (2014).

96. Xiao, W., Warrington, K. H., Hearing, P., Hughes, J. & Muzyczka, N. Adenovirus-facilitated nuclear translocation of adeno-associated virus type 2. *J Virol* **76**, 11505–11517 (2002).

97. Hansen, J., Qing, K., Kwon, H.-J., Mah, C. & Srivastava, A. Impaired intracellular trafficking of adeno-associated virus type 2 vectors limits efficient transduction of murine fibroblasts. *J Virol* **74**, 992–996 (2000).

98. Nonnenmacher, M. E., Cintrat, J.-C., Gillet, D. & Weber, T. Syntaxin 5-dependent retrograde transport to the trans-Golgi network is required for adeno-associated virus transduction. *J Virol* **89**, 1673–1687 (2015).

99. Pajusola, K. *et al.* Cell-type-specific characteristics modulate the transduction efficiency of adeno-associated virus type 2 and restrain infection of endothelial cells. *J Virol* **76**, 11530–11540 (2002).

100. Liu, Y., Joo, K. I. & Wang, P. Endocytic processing of adeno-associated virus type 8 vectors for transduction of target cells. *Gene Ther* **20**, 308–317 (2013).

101. Zhong, L. *et al.* Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc Natl Acad Sci U S A* **105**, 7827–7832 (2008).

102. Li, C. *et al.* Adeno-associated virus capsid antigen presentation is dependent on endosomal escape. *J Clin Invest* **123**, 1390–1401 (2013).

103. Finn, J. D. *et al.* Proteasome inhibitors decrease AAV2 capsid derived peptide epitope presentation on MHC class I following transduction. *Mol Ther* **18**, 135–142 (2010).

104. Johnson, J. S. & Samulski, R. J. Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus. *J Virol* **83**, 2632–2644 (2009).

105. Salganik, M. *et al.* Evidence for pH-dependent protease activity in the adeno-associated virus capsid. *J Virol* **86**, 11877–11885 (2012).

106. Akache, B. *et al.* A two-hybrid screen identifies cathepsins B and L as uncoating factors for adeno-associated virus 2 and 8. *Mol Ther* **15**, 330–339 (2007).

107. Meisen, W. H. *et al.* Pooled Screens Identify GPR108 and TM9SF2 as Host Cell Factors Critical for AAV Transduction. *Mol Ther Methods Clin Dev* **17**, 601–611 (2020).

108. Zhang, R. *et al.* Adeno-associated virus 2 bound to its cellular receptor AAVR. *Nat Microbiol* **4**, 675–682 (2019).

109. Johnson, J. S. *et al.* Mutagenesis of adeno-associated virus type 2 capsid protein VP1 uncovers new roles for basic amino acids in trafficking and cell-specific transduction. *J Virol* **84**, 8888–8902 (2010).

110. Grieger, J. C., Snowdy, S. & Samulski, R. J. Separate basic region motifs within the adeno-associated virus capsid proteins are essential for infectivity and assembly. *J Virol* **80**, 5199–5210 (2006).

111. Nicolson, S. C. & Samulski, R. J. Recombinant adeno-associated virus utilizes host cell nuclear import machinery to enter the nucleus. *J Virol* **88**, 4132–4144 (2014).

112. Kelich, J. M. *et al.* Super-resolution imaging of nuclear import of adeno-associated virus in live cells. *Mol Ther Methods Clin Dev* **2**, 15047 (2015).

113. Horowitz, E. D. *et al.* Biophysical and ultrastructural characterization of adeno-associated virus capsid uncoating and genome release. *J Virol* **87**, 2994–3002 (2013).

114. Johnson, J. S. & Samulski, R. J. Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus. *J Virol* **83**, 2632–2644 (2009).

115. Ferrari, F. K., Samulski, T., Shenk, T. & Samulski, R. J. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J Virol* **70**, 3227–3234 (1996).

116. Fisher, K. J. *et al.* Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *J Virol* **70**, 520–532 (1996).

117. Berns, K. I. & Muzyczka, N. AAV: An Overview of Unanswered Questions. *Hum Gene Ther* **28**, 308–313 (2017).

118. Verdera, H. C., Kuranda, K. & Mingozzi, F. AAV Vector Immunogenicity in Humans: A Long Journey to Successful Gene Transfer. *Molecular Therapy* **28**, 723–746 (2020).

119. Grieger, J. C. & Samulski, R. J. Adeno-associated virus as a gene therapy vector: Vector development, production and clinical applications. *Adv Biochem Eng Biotechnol* **99**, 119–145 (2005).

120. Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nature Protocols 2006 1:3* **1**, 1412–1428 (2006).

121. ZOLGENSMA | FDA. https://www.fda.gov/vaccines-blood-biologics/zolgensma.

122. Dumont, J., Euwart, D., Mei, B., Estes, S. & Kshirsagar, R. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit Rev Biotechnol* **36**, 1110–1122 (2016).

123. Walsh, G. Biopharmaceutical benchmarks 2018. *Nature Biotechnology 2018 36:12* **36**, 1136–1145 (2018).

124. Tan, E., Chin, C. S. H., Lim, Z. F. S. & Ng, S. K. HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. *Front Bioeng Biotechnol* **9**, 796991 (2021).

125. Hong, M. *et al.* Genetic engineering of baculovirus-insect cell system to improve protein production. *Front Bioeng Biotechnol* **10**, 994743 (2022).

126. Palomares, L. A., Srivastava, I. K., Ramírez, O. T. & Cox, M. M. J. Glycobiotechnology of the Insect Cell-Baculovirus Expression System Technology. *Adv Biochem Eng Biotechnol* **175**, 71–92 (2018).

127. Urabe, M., Ding, C. & Kotin, R. M. Insect Cells as a Factory to Produce Adeno-Associated Virus Type 2 Vectors. *https://home.liebertpub.com/hum* **13**, 1935–1943 (2004).

128. Kotin, R. M. Large-scale recombinant adeno-associated virus production. *Hum Mol Genet* **20**, R2–R6 (2011).

129. Hu, Y. C. Baculovirus Vectors for Gene Therapy. *Adv Virus Res* **68**, 287–320 (2006).

130. Monie, A., Hung, C. F., Roden, R. & Wu, T. C. Cervarix™: a vaccine for the prevention of HPV 16, 18-associated cervical cancer. *Biologics* **2**, 107–113 (2008).

131. Heo, Y. A. Etranacogene Dezaparvovec: First Approval. *Drugs* **83**, 347–352 (2023).

132. Quadros, R. M. *et al.* Easi-CRISPR: A robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. *Genome Biol* **18**, 1–15 (2017).

133. Wu, S.-H. S. *et al.* SCON - A Short Conditional intrON for conditional knockout with one-step zygote injection. *bioRxiv* 2021.05.09.443220 (2021) doi:10.1101/2021.05.09.443220.

134. Wu, Z., Yang, H. & Colosi, P. Effect of Genome Size on AAV Vector Packaging. *Molecular Therapy* **18**, 80–86 (2010).

135. Schultz, B. R. & Chamberlain, J. S. Recombinant Adeno-associated Virus Transduction and Integration. *Mol Ther* **16**, 1189 (2008).

136. Chen, S. *et al.* CRISPR-READI: Efficient Generation of Knockin Mice by CRISPR RNP Electroporation and AAV Donor Infection. *Cell Rep* **27**, 3780-3789.e4 (2019).

137. N, M. *et al.* Intra-embryo Gene Cassette Knockin by CRISPR/Cas9-Mediated Genome Editing with Adeno-Associated Viral Vector. *iScience* **9**, 286–297 (2018).

138. Subramanian, M. *et al.* RNAi-mediated rheostat for dynamic control of AAV-delivered transgenes. *Nat Commun* **14**, (2023).

139. Piepho, A. B. *et al.* Micro-dystrophin gene therapy demonstrates long-term cardiac efficacy in a severe Duchenne muscular dystrophy model. *Mol Ther Methods Clin Dev* **28**, 344–354 (2023).

140. Kawabata, H. *et al.* Improving cell-specific recombination using AAV vectors in the murine CNS by capsid and expression cassette optimization. *Mol Ther Methods Clin Dev* **32**, (2024).

141. Liao, H. K. *et al.* In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. *Cell* **171**, 1495-1507.e15 (2017).

142. Wolter, J. M. *et al.* Cas9 gene therapy for Angelman syndrome traps Ube3a-ATS long non-coding RNA. *Nature* **587**, (2020).

143. De Gasparo, R. *et al.* Bispecific IgG neutralizes SARS-CoV-2 variants and prevents escape in mice. *Nature 2021 593:7859* **593**, 424–428 (2021).

144. Israelow, B. *et al.* Mouse model of SARS-CoV-2 reveals inflammatory role of type I interferon signaling. *J Exp Med* **217**, (2020).

145. Kuo, C. C. *et al.* Differential modulation of methamphetamine-mediated behavioral sensitization by overexpression of Mu opioid receptors in nucleus accumbens and ventral tegmental area. *Psychopharmacology (Berl)* **233**, 661–672 (2016).

146. Chen, Z. *et al.* SARS-CoV-2 immunity in animal models. *Cellular & Molecular Immunology 2024 21:2* **21**, 119–133 (2024).

147. Nickl, P., Raishbrook, M. J., Syding, L. A. & Sedlacek, R. Advances in Modelling COVID-19 in Animals. *Frontiers in Drug Discovery* **2**, 899587 (2022).

148. Allocca, M. *et al.* Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. *J Clin Invest* **118**, 1955–1964 (2008).

149. Grieger, J. C. & Samulski, R. J. Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *J Virol* **79**, 9933 (2005).

150. Ibreljic, N., Draper, B. E. & Lawton, C. W. Recombinant AAV genome size effect on viral vector production, purification, and thermostability. *Mol Ther Methods Clin Dev* **32**, 101188 (2024).

151. Halbert, C. L., Allen, J. M. & Miller, D. D. Efficient mouse airway transduction following recombination between AAV vectors carrying parts of a larger gene. *Nat Biotechnol* **20**, 697–701 (2002).

152. Reguera, J., Carreira, A., Riolobos, L., Almendral, J. M. & Mateu, M. G. Role of interfacial amino acid residues in assembly, stability, and conformation of a spherical virus capsid. *Proc Natl Acad Sci U S A* **101**, 2724–2729 (2004).

153. Duan, D. Systemic AAV Micro-dystrophin Gene Therapy for Duchenne Muscular Dystrophy. *Mol Ther* **26**, 2337–2356 (2018).

154. England, S. B. *et al.* Very mild muscular dystrophy associated with the deletion of 46% of dystrophin. *Nature* **343**, 180–182 (1990).

155. Tillotson, R. *et al.* Radically truncated MeCP2 rescues Rett syndrome-like neurological defects. *Nature* **550**, (2017).

156. Zhang, W., Li, L., Su, Q., Gao, G. & Khanna, H. Gene Therapy Using a miniCEP290 Fragment Delays Photoreceptor Degeneration in a Mouse Model of Leber Congenital Amaurosis. *Hum Gene Ther* **29**, 42–50 (2018).

157. Marrone, L., Marchi, P. M. & Azzouz, M. Circumventing the packaging limit of AAV-mediated gene replacement therapy for neurological disorders. *Expert Opin Biol Ther* **22**, 1163–1176 (2022).

158. Zhang, Y. & Duan, D. Novel mini-dystrophin gene dual adeno-associated virus vectors restore neuronal nitric oxide synthase expression at the sarcolemma. *Hum Gene Ther* **23**, 98–103 (2012).

159. Odom, G. L., Gregorevic, P., Allen, J. M. & Chamberlain, J. S. Gene Therapy of mdx Mice With Large Truncated Dystrophins Generated by Recombination Using rAAV6. *Molecular Therapy* **19**, 36–45 (2011).

160. Pryadkina, M. *et al.* A comparison of AAV strategies distinguishes overlapping vectors for efficient systemic delivery of the 6.2 kb Dysferlin coding sequence. *Mol Ther Methods Clin Dev* **2**, 15009 (2015).

161. Sondergaard, P. C. *et al.* AAV. Dysferlin overlap vectors restore function in dysferlinopathy animal models. *Ann Clin Transl Neurol* **2**, 256–270 (2015).

162. Trapani, I. *et al.* Effective delivery of large genes to the retina by dual AAV vectors. *EMBO Mol Med* **6**, 194–211 (2014).

163. Liang, F., Han, M., Romanienko, P. J. & Jasin, M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc Natl Acad Sci U S A* **95**, 5172–5177 (1998).

164. Duan, D., Yue, Y., Yan, Z. & Engelhardt, J. F. A new dual-vector approach to enhance recombinant adeno-associated virus-mediated gene expression through intermolecular cis activation. *Nature Medicine 2000 6:5* **6**, 595–598 (2000).

165. Nakai, H., Storm, T. A. & Kay, M. A. Increasing the size of rAAV-mediated expression cassettes in vivo by intermolecular joining of two complementary vectors. *Nat Biotechnol* **18**, 527–532 (2000).

166. Sun, L., Li, J. & Xiao, X. Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization. *Nature Medicine 2000 6:5* **6**, 599–602 (2000).

167. Shah, N. H. & Muir, T. W. Inteins: nature's gift to protein chemists. *Chem Sci* **5**, 446–461 (2013).

168. Tornabene, P. *et al.* Intein-mediated protein trans-splicing expands adeno-associated virus transfer capacity in the retina. *Sci Transl Med* **11**, (2019).

169. Li, J., Sun, W., Wang, B., Xiao, X. & Liu, X. Q. Protein trans-splicing as a means for viral vector-mediated in vivo gene therapy. *Hum Gene Ther* **19**, 958–964 (2008).

170. Tiffany, M. & Kay, M. A. 253. Expanded Packaging Capacity of AAV by Lumenal Charge Alteration. *Molecular Therapy* **24**, S99–S100 (2016).

171. Zolotukhin, S. *et al.* *Recombinant Adeno-Associated Virus Purification Using Novel Methods Improves Infectious Titer and Yield*. *Gene Therapy* vol. 6 (1999).

172. Khanal, O., Kumar, V. & Jin, M. Adeno-associated viral capsid stability on anion exchange chromatography column and its impact on empty and full capsid separation. *Mol Ther Methods Clin Dev* **31**, (2023).

173. Florea, M. *et al.* High-efficiency purification of divergent AAV serotypes using AAVX affinity chromatography. *Mol Ther Methods Clin Dev* **28**, 146 (2023).

174. Crosson, S. M., Dib, P., Smith, J. K. & Zolotukhin, S. Helper-free Production of Laboratory Grade AAV and Purification by Iodixanol Density Gradient Centrifugation. *Mol Ther Methods Clin Dev* **10**, 1–7 (2018).

175. Lam, A. K. *et al.* Comprehensive Comparison of AAV Purification Methods: Iodixanol Gradient Centrifugation vs. Immuno-Affinity Chromatography. *Adv Cell Gene Ther* **2023**, 1–12 (2023).

176. Maguire, C. A. *et al.* Microvesicle-associated AAV vector as a novel gene delivery system. *Molecular Therapy* **20**, 960–971 (2012).

177. Deshetty, U. M., Sil, S. & Buch, S. Gene therapy for the heart: encapsulated viruses to the rescue. *Extracell Vesicles Circ Nucleic Acids 2024;5:114-8.* **5**, 114–118 (2024).

178. György, B. & Maguire, C. A. Extracellular vesicles: nature's nanoparticles for improving gene transfer with adeno-associated virus vectors. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* **10**, e1488 (2018).

179. Wang, B. Z., Luo, L. J. & Vunjak-Novakovic, G. RNA and Protein Delivery by Cell-Secreted and Bioengineered Extracellular Vesicles. *Adv Healthc Mater* **11**, e2101557 (2022).

180. Wang, Y. *et al.* Decoding cellular mechanism of recombinant adeno-associated virus (rAAV) and engineering host-cell factories toward intensified viral vector manufacturing. *Biotechnol Adv* **71**, 108322 (2024).

181. Rufino-Ramos, D. *et al.* Extracellular vesicles: Novel promising delivery systems for therapy of brain diseases. (2017) doi:10.1016/j.jconrel.2017.07.001.

182. Théry, C., Amigorena, S., Raposo, G. & Clayton, A. Isolation and Characterization of Exosomes from Cell Culture Supernatants and Biological Fluids. *Curr Protoc Cell Biol* **30**, (2006).

183. Abels, E. R. & Breakefield, X. O. Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake. doi:10.1007/s10571-016-0366-z.

184. Kalra, H., Drummen, G. P. C. & Mathivanan, S. Focus on Extracellular Vesicles: Introducing the Next Small Big Thing. *International Journal of Molecular Sciences 2016, Vol. 17, Page 170* **17**, 170 (2016).

185. Zhang, D. X. *et al.* The Biology and Therapeutic Applications of Red Blood Cell Extracellular Vesicles. *Erythrocyte* (2018) doi:10.5772/INTECHOPEN.81758.

186. Schiller, L. T., Lemus-Diaz, N., Ferreira, R. R., Böker, K. O. & Gruber, J. Enhanced Production of Exosome-Associated AAV by Overexpression of the Tetraspanin CD9. *Mol Ther Methods Clin Dev* **9**, 278–287 (2018).

187. György, B., Fitzpatrick, Z., Crommentuijn, M. H. W., Mu, D. & Maguire, C. A. Naturally enveloped AAV vectors for shielding neutralizing antibodies and robust gene delivery in vivo. *Biomaterials* **35**, 7598–7609 (2014).

188. Zolotukhin, S. *et al.* Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther* **6**, 973–985 (1999).

189. Wright, J. F. *et al.* Identification of factors that contribute to recombinant AAV2 particle aggregation and methods to prevent its occurrence during vector purification and formulation. *Mol Ther* **12**, 171–178 (2005).

190. Raposo, G. & Stoorvogel, W. Extracellular vesicles: exosomes, microvesicles, and friends. *J Cell Biol* **200**, 373–383 (2013).

191. Abels, E. R. & Breakefield, X. O. Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake. *Cell Mol Neurobiol* **36**, 301 (2016).

192. Benskey, M. J., Sandoval, I. M. & Manfredsson, F. P. Continuous collection of adeno-associated virus from producer cell medium significantly increases total viral yield. *Hum Gene Ther Methods* **27**, 32–45 (2016).

193. Rider, M. A., Hurwitz, S. N. & Meckes, D. G. ExtraPEG: A polyethylene glycol-based method for enrichment of extracellular vesicles. *Sci Rep* **6**, (2016).

194. Schiller, L. T., Lemus-Diaz, N., Ferreira, R. R., Böker, K. O. & Gruber, J. Enhanced Production of Exosome-Associated AAV by Overexpression of the Tetraspanin CD9. (2018) doi:10.1016/j.omtm.2018.03.008.

195. Lai, C. P. *et al.* Dynamic Biodistribution of Extracellular Vesicles In Vivo Using a Multimodal Imaging Reporter. *ACS Nano* **8**, 483 (2014).

196. Aurnhammer, C. *et al.* Universal real-time PCR for the detection and quantification of adeno-associated virus serotype 2-derived inverted terminal repeat sequences. *Hum Gene Ther Methods* **23**, 18–28 (2012).

197. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res* **29**, e45 (2001).

198.   Théry, C. *et al.* Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J Extracell Vesicles* **7**, (2018).

199.   Hudry, E. *et al.* Exosome-associated AAV vector as a robust and convenient neuroscience tool. *Gene Ther* **23**, 380–392 (2016).

200.   Rosenkrans, Z. T. *et al.* Investigating the In Vivo Biodistribution of Extracellular Vesicles Isolated from Various Human Cell Sources Using Positron Emission Tomography. *Mol Pharm* **21**, 4324–4335 (2024).

201.   Zincarelli, C., Soltys, S., Rengo, G. & Rabinowitz, J. E. Analysis of AAV Serotypes 1–9 Mediated Gene Expression and Tropism in Mice After Systemic Injection. *Molecular Therapy* **16**, 1073–1080 (2008).

202.   Görgens, A. *et al.* Identification of storage conditions stabilizing extracellular vesicles preparations. *J Extracell Vesicles* **11**, (2022).

203.   Brooks, A. R. *et al.* Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. *J Gene Med* **6**, 395–404 (2004).

204.   Dutheil, N. *et al.* Characterization of the Mouse Adeno-Associated Virus AAVS1 Ortholog. *J Virol* **78**, 8917 (2004).

205.   Chandler, R. J., Sands, M. S. & Venditti, C. P. Recombinant Adeno-Associated Viral Integration and Genotoxicity: Insights from Animal Models. *Hum Gene Ther* **28**, 314 (2017).

206.   Scialo, F. *et al.* ACE2: The Major Cell Entry Receptor for SARS-CoV-2. *Lung* **198**, 867 (2020).

207.   Winkler, E. S. *et al.* SARS-CoV-2 infection of human ACE2-transgenic mice causes severe lung inflammation and impaired function. *Nat Immunol* **21**, 1327–1335 (2020).

208.   Steiner, S. *et al.* SARS-CoV-2 biology and host interactions. *Nature Reviews Microbiology 2024 22:4* **22**, 206–225 (2024).

209.   Hassan, A. O. *et al.* A SARS-CoV-2 Infection Model in Mice Demonstrates Protection by Neutralizing Antibodies. *Cell* **182**, 744 (2020).

210.   Rawle, D. J. *et al.* ACE2-lentiviral transduction enables mouse SARS-CoV-2 infection and mapping of receptor interactions. *PLoS Pathog* **17**, e1009723 (2021).

211.   Katzman, C. *et al.* Modeling SARS-CoV-2 Infection in Mice Using Lentiviral hACE2 Vectors Infers Two Modes of Immune Responses to SARS-CoV-2 Infection. *Viruses* **14**, (2021).

212. Tseng, C.-T. K. *et al.* Severe Acute Respiratory Syndrome Coronavirus Infection of Mice Transgenic for the Human Angiotensin-Converting Enzyme 2 Virus Receptor. *J Virol* **81**, 1162–1173 (2007).

213. Bruter, A. V. *et al.* Novel transgenic mice with Cre-dependent co-expression of GFP and human ACE2: a safe tool for study of COVID-19 pathogenesis. *Transgenic Res* **30**, 289–301 (2021).

214. McCray, P. B. J. *et al.* Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. *J Virol* **81**, 813–821 (2007).

215. Jiang, R. Di *et al.* Pathogenesis of SARS-CoV-2 in Transgenic Mice Expressing Human Angiotensin-Converting Enzyme 2. *Cell* **182**, 50-58.e8 (2020).

216. Bao, L. *et al.* The pathogenicity of SARS-CoV-2 in hACE2 transgenic mice. *Nature 2020 583:7818* **583**, 830–833 (2020).

217. Zhang, W. *et al.* Mouse genome rewriting and tailoring of three important disease loci. *Nature* **623**, 423 (2023).

218. Williamson, E. J. *et al.* Factors associated with COVID-19-related death using OpenSAFELY. *Nature* **584**, 430–436 (2020).

219. Zhang, Y. N. *et al.* Increased morbidity of obese mice infected with mouse-adapted SARS-CoV-2. *Cell Discov* **7**, (2021).

220. Jiang, X. *et al.* The Effects of AT1R Blocker on the Severity of COVID-19 in Hypertensive Inpatients and Virulence of SARS-CoV-2 in Hypertensive hACE2 Transgenic Mice. *J Cardiovasc Transl Res* **15**, 38–48 (2022).

221. Ebinge, J. E. *et al.* Pre-existing traits associated with Covid-19 illness severity. *PLoS One* **15**, (2020).

222. Nickl, P. *et al.* Multistep allelic conversion in mouse pre-implantation embryos by AAV vectors. *Scientific Reports 2024 14:1* **14**, 1–13 (2024).

223. Grainge, I. & Jayaram, M. The integrase family of recombinases: organization and function of the active site. *Mol Microbiol* **33**, 449–456 (1999).

224. Stelzer, Y., Shivalila, S., Soldner, F., Markoulaki, S. & Jaenisch, R. *Tracing Dynamic Changes of DNA Methylation at Single-Cell Resolution*.

225. György, B., Fitzpatrick, Z., Crommentuijn, M. H. W., Mu, D. & Maguire, C. A. Naturally enveloped AAV vectors for shielding neutralizing antibodies and robust gene delivery invivo. *Biomaterials* **35**, 7598–7609 (2014).

226. Maguire, C. A. *et al.* Microvesicle-associated AAV vector as a novel gene delivery system. *Molecular Therapy* **20**, 960–971 (2012).

227. Schiller, L. T., Lemus-Diaz, N., Ferreira, R. R., Böker, K. O. & Gruber, J. Enhanced Production of Exosome-Associated AAV by Overexpression of the Tetraspanin CD9. *Mol Ther Methods Clin Dev* **9**, 278–287 (2018).

228. Murayama, E. *et al.* Tracing hematopoietic precursor migration to successive hematopoietic organs during zebrafish development. *Immunity* **25**, 963–75 (2006).

229. Li, X. *et al.* Extracellular Vesicle-Encapsulated Adeno-Associated Viruses for Therapeutic Gene Delivery to the Heart. *Circulation* **148**, 405–425 (2023).

230. Vyas, P., Balakier, H. & Librach, C. L. Ultrastructural identification of CD9 positive extracellular vesicles released from human embryos and transported through the zona pellucida. *Syst Biol Reprod Med* **65**, 273–280 (2019).

231. Machtinger, R., Laurent, L. C. & Baccarelli, A. A. Extracellular vesicles: roles in gamete maturation, fertilization and embryo implantation. *Hum Reprod Update* **22**, 182 (2016).

8. Publication and author contributions

<u>Doctoral thesis contains the following publications:</u>

Nickl, P., Jenickova, I., Elias, J., Kasparek, P., Barinka, C., Kopkanova, J., & Sedlacek, R. (2024). **Multistep allelic conversion in mouse pre-implantation embryos by AAV vectors**. *Scientific Reports 2024 14:1*, *14*(1), 1–13. https://doi.org/10.1038/s41598-024-70853-1

- Petr Nickl conceived the experiments, conducted cloning work, produced AAVs, planned and performed animal work, and analyzed the results. He contributed to writing, reviewing, and editing the manuscript.

de Gasparo, R., Pedotti, M., Simonelli, L., **Nickl, P**., Muecksch, F., Cassaniti, I., Percivalle, E., Lorenzi, J. C. C., Mazzola, F., Magrì, D., Michalcikova, T., Haviernik, J., Honig, V., Mrazkova, B., Polakova, N., Fortova, A., Tureckova, J., Iatsiuk, V., di Girolamo, S., … Varani, L. (2021). **Bispecific IgG neutralizes SARS-CoV-2 variants and prevents escape in mice.** *Nature 2021 593:7859*, *593*(7859), 424–428. https://doi.org/10.1038/s41586-021-03461-y

- Petr Nickl performed mouse experiments, analyzed the results, and conceived and designed the mouse model.

Nickl, P., Raishbrook, M. J., Syding, L. A., & Sedlacek, R. (2022). **Advances in Modelling COVID-19 in Animals.** *Frontiers in Drug Discovery*, *0*, 8. https://doi.org/10.3389/FDDSV.2022.899587

- PN, MR, and LS wrote the article with equal contributions.

<u>Manuscript in preparation:</u>

Nickl P., Barbiera M., Zini J., Nickl T., Ushiki A., Ahituv N., Prochazka J., Yliperttula M., Sedlacek R. **Scalable Production of rAAV Vectors via Extracellular Vesicle-Mediated Purification for Gene Therapy and Transgenesis**

- Petr Nickl concieved the experiment, conducted cloning work, produced EV-AAV and AAV, planned and performed animal work, and analyzed the result. He contributes to writing, reviewing, and editing the manuscript.

PD. Dr. rer. nat. habil. Radislav Sedláček
supervisor