Heliyon 11 (2025) e43732

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Improved production and quality of lentiviral vectors by major-splice-donor mutation and co-expression of a novel U1 snRNA-based enhancer

J. Wright [a], B.M. Alberts [a], A.J.M. Hood [a,b], C. Nogueira [a,c], Z. Miskolczi [a,d,1], C.R. Vieira [a,2], D. Chipchase [a], C.M. Lamont [a], O. Goodyear [a], L.J. Moyce [a], M. Soyombo [a,e], D. Blount [a,f,3], A.L. Keating [a,g], T. Coradin [a,h], H. Huang [a], M. Martin-Urdiroz [a], S. Ferluga [a,i], K.A. Mitrophanous [a], N.G. Clarkson [a], D.C. Farley [a,*]

[a] *Oxford Biomedica (UK) Ltd, Windrush Court, Transport Way, Oxford, OX4 6LT, UK*
[b] *Division of Structural Biology, Nuffield Department of Medicine, University of Oxford, The Centre for Human Genetics, Roosevelt Drive, Oxford, OX3 7BN, UK*
[c] *Immunocore Limited, 92 Park Drive, Milton Park, Abingdon, Oxon, OX14 4RY, UK*
[d] *Adaptimmune Limited, 60 Jubilee Avenue, Milton Park, Oxfordshire, Abingdon, OX14 4RX, UK*
[e] *bioMérieux, Chineham Gate, Crockford Ln, Chineham, Basingstoke, RG24 8NA, UK*
[f] *Barinthus Biotherapeutics, Units 6–10, Zeus Building, Rutherford Ave, Harwell, Didcot, OX11 0DF, UK*
[g] *Department of Oncology, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford, OX3 7DQ, UK*
[h] *iosBio Ltd, Hayworthe House, Market Square, Haywards Heath, RH161DB, UK*
[i] *Kings College London Gene Therapy Vector Facility (GTVF), 123 Coldharbour Lane, London, SE5 9NU, UK*

A B S T R A C T

The clinical success in using HIV-1 based Lentiviral vectors (LVs) to deliver therapeutic genes is built upon the safety features of 3rd generation LVs. However, the Cell and Gene Therapy industry as well as regulatory agencies are continually seeking improvements in product quality and safety. LV RNA genomes retain complex HIV-1 *cis*-acting sequences that govern production of unspliced/spliced vector (v)RNA, including the major splice donor (MSD) within the packaging signal. Substantial aberrant splicing from the MSD to strong/cryptic splice acceptors within transgene cassettes can occur during LV production. Such aberrant vRNAs encode the transgene, explaining why tissue-specific promoter-driven transgenes appear to be leaky during LV production. Additionally, the TRiP system™ (Transgene Repression In vector Production), which represses translation of transgenes during production to enhance titres and/or omit the transgene protein from product, is less efficacious on these spliced vRNAs. We show that resultant spliced forms of vRNA can be detected in LV product and can be converted to episomal cDNA in target cells. MSD-mutation reduces LV titre in tat-independent systems, and therefore these limitations remain unsolved in contemporary 3rd generation LV platforms. We present a novel LV

* Corresponding author.
E-mail addresses: j.wright@oxb.com (J. Wright), b.alberts@oxb.com (B.M. Alberts), alasdair.hood@strubi.ox.ac.uk (A.J.M. Hood), Cristina.nogueira@immunocore.com (C. Nogueira), zsofia.miskolczi@outlook.com (Z. Miskolczi), clrvv22@gmail.com (C.R. Vieira), dandchipchase@googlemail.com (D. Chipchase), c.lamont@oxb.com (C.M. Lamont), o.goodyear@oxb.com (O. Goodyear), l.moyce@oxb.com (L.J. Moyce), mayowa.s8@gmail.com (M. Soyombo), Daniel.blount@matchbiotx.com (D. Blount), amy.keating@queens.ox.ac.uk (A.L. Keating), tizianacoradin@libero.it (T. Coradin), hlhuang318@hotmail.com (H. Huang), M.MUrdiroz@oxb.com (M. Martin-Urdiroz), sferluga75@gmail.com (S. Ferluga), k.mitrophanous@oxb.com (K.A. Mitrophanous), n.clarkson@oxb.com (N.G. Clarkson), d.farley@oxb.com (D.C. Farley).
[1] Current address: Turbine, Budapest, Szigony street 26-32, 1083, Hungary.
[2] Current address: Indigo Medical, Second floor, Vega House, Eastlake Park, Fox Milne, Milton Keynes, MK15 0DF, UK.
[3] Current address: MatchBio Ltd., Africa House, 70 Kingsway, London, WC2B 6AH, UK.

https://doi.org/10.1016/j.heliyon.2025.e43732
Received 5 August 2024; Received in revised form 8 August 2025; Accepted 11 August 2025
Available online 4 September 2025

platform – referred to as the TetraVecta™ System - utilising optimised MSD-inactivating sequences, along with a new class of vRNA enhancer based on modified U1 snRNA. These enhancers, co-expressed during production, rescue the output titres of MSD-mutated LVs. MSD-mutated LVs do not generate spliced forms of vRNAs or the resultant episomal cDNAs, work optimally with the TRiP system™, and represent a further development in the utility and safety of LVs.

1. Introduction

The use of lentiviral vectors (LVs) is an established and efficient means to deliver a large therapeutic payload to dividing or non-dividing cells to achieve long term gene expression [1]. One advantage of LVs is the protection of genetic cargo within their capsids, both spatially and temporally until close to the point of integration into the host cell DNA, in a well characterised and orchestrated manner [2]. Conversely, DNA based vectors or donor DNA that remain episomal, may be manipulated at low frequency in the host cell, such as by the DNA repair machinery [3–7]. The development of LVs for *in vivo* indications is being re-invigorated through envelope re-targeting [8,9], and improvements in production scales [10]. In parallel, the regulatory authorities are placing much more importance on the safety/quality aspect of viral vector gene therapy products generally [11]. Therefore, as the industry continues to

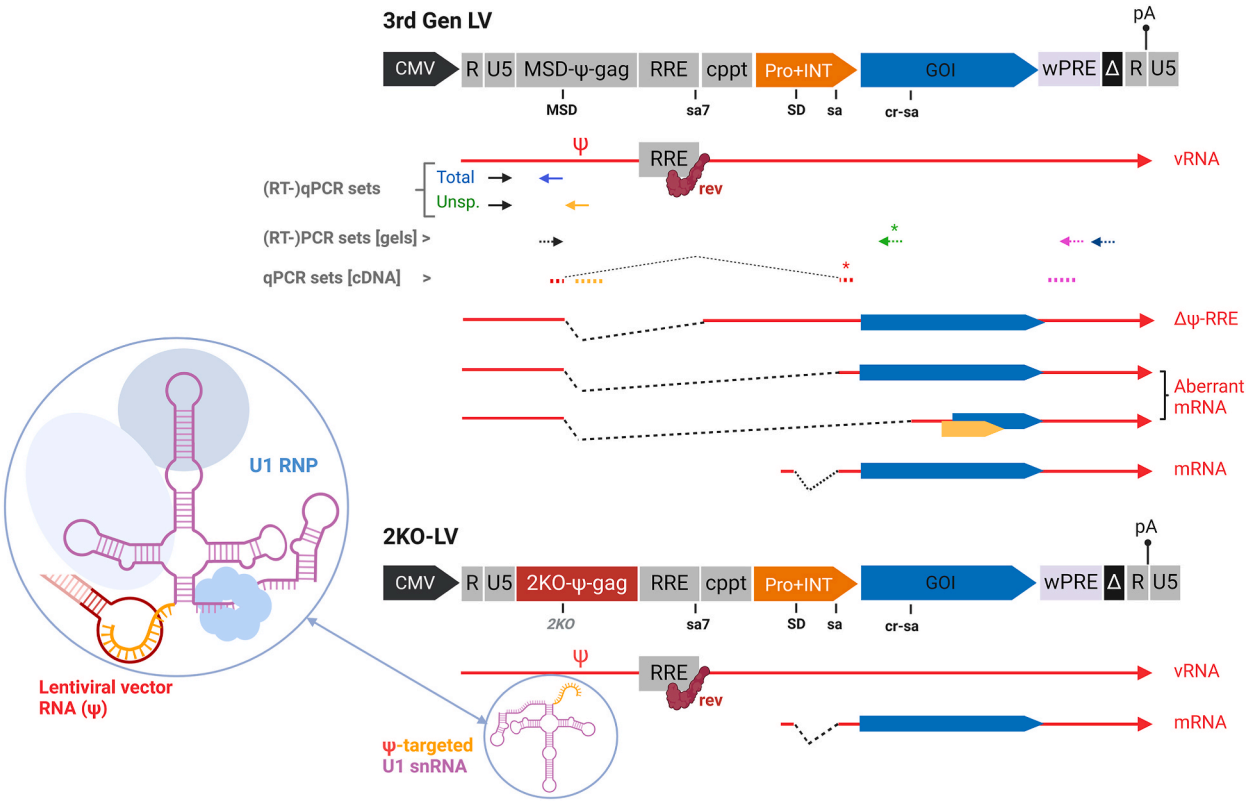


Fig. 1. Overview of 3rd generation and '2KO' LV genome expression cassettes and their RNA products. The salient features of lentiviral vector genomes, with the external promoter (Cytomegalovirus promoter [CMV]) and polyadenylation site (pA) positions defining the transcription unit. The 5′ and 3′ LTR sequences (R-U5), packaging signal (ψ-gag), Rev-response element (RRE), central polypurine tract (cppt), tRNA primer binding site (PBS) and 3′ppt (not displayed) are necessary cis-elements. The internal transgene expression cassette is driven by a promoter, sometimes containing an intron (Pro + INT), and encodes a gene of interest (GOI). Typically, LVs utilise a post-transcriptional regulatory element (PRE) to enhance GOI expression in target cells, such as the Woodchuck Hepatitis Virus (w)PRE. Self-inactivating (SIN)-LTRs are deleted for the native HIV-1 U3 promoter/enhancer (Δ). The RRE is accompanied with splice acceptor (sa) 7 from HIV-1 to form an artificial intron from the MSD-to-sa7, splicing out of which may lead to 'Δψ-RRE'. Binding of Rev to the RRE enables full length, unspliced vRNA to appear in the cytoplasm. However, 'aberrant' splicing from the MSD to strong or cryptic splice acceptor (cr-sa) sites within the transgene cassette can occur even in the presence of Rev. Spliced mRNA encoding the GOI or truncated in-/out-of-frame proteins may be translated during production. 2KO-LV genome expression cassettes are MSD-inactivated by the '2KO' feature. Maximal 2KO-LV output titres are dependent on the co-expression of a modified U1 snRNA that can bind to a target sequence within ψ. The position of primers and probe for (RT)-qPCR assays (SYBR) to assess the relative amount of unspliced and spliced vRNA, by measuring total (blue arrow) or unspliced (orange arrow) vRNA with common forward primer (black). Use of forward (dotted black) and reverse (GFP*, where this was the GOI; dotted green, wPRE; dotted pink, 3′ ppt; dotted navy blue) primers enabled visualisation on a gel of the degree and variety of aberrant mRNAs produced in a single RT-PCR assay. The qPCR primer/probe (Taqman) set positions for integration/cDNA analysis are denoted by the orange (Psi/ψ), pink (wPRE) and red (MSD-to-EF1a splice acceptor* junction) dotted lines (Created in BioRender.com).

seek improvements in manufacturing yields and product quality, attention remains on engineering the genetic components of LVs.

The principal *cis*-acting sequences contained within the HIV-1 vector genomic RNA (vRNA) include: the 5′ packaging sequence (Psi/ψ); the centrally positioned [12] Rev-response element (RRE), on which Rev acts to provide unspliced genomic RNA; and sequences necessary for both reverse transcription and integration steps (see Fig. 1). Two important safety aspects of 3rd generation LVs are the production of vRNA independently of the HIV-1 auxiliary protein Tat (by use of a powerful heterologous constitutive promoter), and the self-inactivating (SIN) 3' LTR, which is devoid of native promoter activity in target cells. Tat has been shown to impact on many cellular functions [13], and its removal from the 3rd generation LV platform is an important safety feature. The complete separation of LV genome, Gag/Pol, Rev and envelope genes to independent expression cassettes also further reduces likelihood of replication-competent lentivirus (RCL) formation during production [14,15].

Output titres of LVs are strongly correlated to the steady-state pool of unspliced, packageable vRNA in the production cell [16]. The Rev-RRE paradigm employed by LVs enables the production of vector RNA harbouring intron-containing transgene cassettes, such as the EF1α promoter-intron. This is functionally analogous to the mechanism HIV-1 uses with Rev-RRE to generate unspliced genomic RNA in the cytoplasm for packaging. However, it is clear this is an imperfect system, as certain introns cannot be retained within LV genomes, requiring inversion of the transgene cassette [17,18]. Further, here we demonstrate that the Rev-RRE configuration is highly variable in its ability to inhibit splicing from the MSD (embedded within ψ) to transgene sequences downstream of the RRE, leading to the production of MSD-spliced vRNAs that can contain transgene coding sequences yet lack the internal promoter. This effect can be found within the literature, although often not the focus of the authors' work, and generally not quantified [19–22].

The 'Transgene Repression In vector Production' (TRiP) system™ utilises a translation-blocking mechanism during LV production to minimise the problematic effects of certain transgene proteins on LV output titres and/or processing [23,24]. We find that these MSD-derived, spliced vRNAs typically encode and express transgene protein during production and can be poorly repressed in the TRiP system™. Depending on abundance, they can also be detected in LV product and converted into episomal cDNA in target cells. Therefore, the purpose of this work was to ablate MSD activity to: (i) simplify the vRNA 'profile' of LV product by increasing the proportion of full length vRNA (quality); (ii) negate any concerns of 'packaged' spliced vRNAs and their potential interference/role with cDNA formation in target cells (quality/performance); and (iii) provide a means of achieving maximal TRiP repression (utility).

We find that mutation of the MSD within 3rd generation LVs leads to a reduction in output titres, in line with others' work [20]. To mitigate this, we employed specific MSD-mutating sequences to generate '2KO-LV' genomes, and a modified U1 snRNA (re-targeted towards the vRNA packaging sequence), which is co-expressed with LV components during production. These novel U1 snRNA-like enhancers increase the steady-state pool of vRNA during production through a mechanism independent of polyadenylation suppression, the currently understood method of action of U1 snRNA 'telescripting' utilised in U1-interference [25,26]. The TRiP system™ is optimally employed within 2KO-LVs, and the lack of spliced vRNA detected in 2KO-LV product correlated with the absence of related cDNAs in target cells.

2. Materials & methods

2.1. Plasmids

HIV-1 lentiviral vector genomes and packaging plasmids were constructed as described previously [58–60]. Mutations to the MSD were made by gene synthesis and subsequent cloning using *Nde*I and *Psp*XI restriction sites. The U1 expression cassettes were based on the endogenous human U1 gene (GenBank J00318.1; nucleotides 41–724). This entire U1 sequence including the endogenous promoter and termination signals, but excluding the native splice donor complementary sequence, was placed into pMA (Thermo Fisher Scientific) to generate pUWonder-*Sap*I. Into pUWonder-*Sap*I all re-targeting sequence variants for the panel screens were inserted via a ready-made double-stranded oligo into *Sap*I-digested backbone, ensuring the 'AT' dinucleotide was maintained at the transcription start site. Targeting sequences are reported in Supplementary Table S1. The post-panel screen variant of 256U1[15 nt] was cloned into a re-derived KanR backbone, called pRKH-256U1, and used for scale-up studies. pEF1a-TRAP expression cassettes were constructed as described previously [23]. The pCMV-ψ-MSD|EF1aSA-(± tbsX)-GFP expression cassettes were constructed using gene synthesis of the cDNA encompassing the splice junction of the MSD into the EF1a splice acceptor. These fragments were then cloned into a lentiviral vector construct, leading to the removal of Psi, gag, RRE, cPPT and EF1α promoter sequences up to the EF1α splice acceptor, but leaving the GFP transgene and both 5′ and 3′ LTRs intact. Fragments were synthesized with either the (kzk)tbsV0 or non-overlapping Kozak tbs sequences upstream of the GFP cassette. The optimal TRiP 5′ UTR-tbsV0 leader sequence is: 5′-CTT TTT CGC AAC **GAG TTT AGC GGA GTG GAG AAG AGC GGA GCC GAG CCT AGC AGA GAC GAG CCG AGA TG**-3′, where the bold sequence is the 55 nt $[KAGNN]_{11}$ tbs consensus, and the underlined sequence is the Kozak sequence. pSV40-Tat was constructed by insertion of 101 amino acid version of Tat (GenBank UTI68367.1) into a pMX vector (Thermo Fisher Scientific) flanked by SV40 promoter and polyadenylation signals. Polyadenylation reporter plasmids were constructed through insertion of GFP-SINLTR-IRES-Gaussia Luciferase into a pSF-EF1a-GFP expression plasmid (OXGENE).

2.2. Cell lines

Vector production was performed in adherent HEK293T cells (ATCC; LGC Standards, UK) or in-house suspension-adapted 'HEK293T-1.65s' cells, which were derived from an adherent HEK293T cell GMP bank. Adherent HEK293T cell lines have been assessed by PCR single-locus technology (Eurofins Medigenomix Forensik GmbH) and verified against the DSMZ database. Adherent HEK293T cells were maintained in Dulbecco's Modified Eagle Media (Sigma) supplemented with 2 mM L-glutamine (Sigma), 10 % (v/

v) Non-essential amino acids (Sigma) and 10 % foetal bovine serum (Gibco). Adherent cells were kept at 37°C in 5 % CO_2 in a static incubator. Suspension HEK293T cells were maintained in Freestyle-293 serum-free media (Thermo Fisher Scientific) supplemented with 0.1 % cholesterol, and kept at 37°C in 5 % CO_2 whilst shaking at 200 r.p.m. Stable TRAP-expressing cells (HEK293T.Ntrp10; referred to as 'HEK293T-TRiP' herein) were developed from HEK293T-1.65s cells using stable transfection of the EF1a-TRAP cassette containing a puromycin resistance marker to enable subsequent selection. All cell lines have been verified as mycoplasma-free.

2.3. Small-scale lentiviral vector production

The standard scale production of lentiviral vectors in adherent HEK293T cells was in 10 cm dishes under the following conditions (all conditions were scaled by area when performed in other formats): HEK293T cells were seeded at 3.5×10^6 cell/plate in 10 mL complete HEK293T media (DMEM (Sigma) containing 10 % heat-inactivated FBS (Gibco), 2 mM L-glutamine (Sigma), 1x NEAA (Sigma)) and incubated at 37°C in 5 % CO2 through-out production. Approximately 24 h later, cells were transfected using the following mass ratios of plasmids: 4.5 μg lentiviral vector genome, 1.4 μg pGagPol, 1.1 μg pRev, 0.7 μg pVSVG. Transfection was mediated by mixing DNA with Lipofectamine® 2000CD in Opti-MEM® according to manufacturer's protocol (Thermo Fisher Scientific). 16–20h post transfection, sodium butyrate was added to the culture to a final concentration of 10 mM. 6h later media was removed from the plate and replaced with fresh, complete DMEM. A single harvest was performed 24h later where supernatants were filtered (0.22 μm) and frozen at −80°C.

The production of lentiviral vectors in suspension HEK293T cells was performed using the following mass ratios of plasmid per 1.6×10^6 cells in 1 mL of Freestyle-293 media: 0.95 μg lentiviral vector genome, 0.1 μg pGagPol, 0.07 μg pVSVG and 0.06 μg pRev. Where appropriate, 0.14 μg of pU1 or pTat plasmid was included. pBluescript was used as a stuffer DNA to ensure equal DNA amounts per transfection where required. Transfection was mediated by mixing cDNA with Lipofectamine® 2000CD in Freestyle media according to the manufacturer's protocol. 16–20h later, sodium butyrate (Sigma-Aldrich, Merck) was added to a final concentration of 10 mM then cultures returned to the incubator. A single harvest was performed 24h later where supernatants were filtered (0.22 μm) before being stored at −80°C.

2.4. Bioreactor-scale lentiviral vector production

HIV-1 lentiviral vectors were generated following multi-plasmid co-transfection of the suspension HEK293T cells. Briefly, cells were inoculated at approximately 1×10^6 cells/mL in serum-free FreeStyle 293 Expression Media (Thermo Fisher Scientific) in glass stirred tank bioreactors (STRs) (Applikon) and agitated using an impellor stirring rate of 290 rpm. Cells were incubated at a temperature of 37 °C, a pH set point of 7.2 and dissolved oxygen was maintained in excess of 20 % throughout using an air/oxygen mix supplied via a sintered bead porous sparger. LV production was instigated via transient co-transfection of cells with vector production plasmids and Lipofectamine as described above. Approximately 24 h prior to vector harvest, sodium butyrate was added as described above. At the termination of the production phase, bioreactor contents were clarified using a 0.2 μm normal flow filter (Pall Corporation) operating at 175 mL/min. Clarified vector was stored at −80 °C prior to use in downstream purification steps.

2.5. Anion-exchange chromatography

For 'technical' grade purification, Lentiviral vector was purified from 250 mL of clarified harvest using Mustang Q XT Acrodiscs (Cytiva). Two tris-based buffers were used, a low-salt buffer A (20 mM Tris, 150 mM NaCl, 2 mM MgCl2, pH 7.2) and a high-salt buffer B (20 mM Tris, 1200 mM NaCl, 2 mM MgCl2, pH 7.2). After sanitizing with 1M NaOH, the membrane was equilibrated with buffer A, washed in two steps with 15 % and 60 % buffer B, and the vector was eluted at 100 % buffer B. The eluted vector was then diluted to 300 mM NaCl and treated with a total of 1000 U of SAN-HQ (Salt Active Nuclease High Quality - ArcticZymes Technologies ASA) at RT for 1 h. After sterile filtration on 0.22 μm syringe filter, the vector was further concentrated by ultracentrifugation (21000×*g*) for 1h at 4 °C, resuspended in the final formulation buffer, aliquoted and stored at −80 °C.

For 'GMP-like'/small downstream model (SDM) grade purification AIEX was conducted using an ÄKTA Avant 150 (Cytiva) and Sartobind® Q nano 1 mL membrane adsorber (Sartorius) in a methodology equivalent to that reported by Pamenter et al. [61]. Following AIEX chromatography, vector was subsequently concentrated and transferred into a final formulation buffer (20 mM Tris, 1 % w/v sucrose, 100 mM sodium chloride and 1 % w/v mannitol) via tangential flow filtration/diafiltration using a polysulfone hollow fibre membrane with a MWCO of 500 kDa (Cytiva).

2.6. Vector transduction and titration by flow cytometry

Ninety-six well plates were seeded with 1.2×10^4 HEK293T cells/well in 100 μL DMEM. 24h later medium was removed and replaced with 50μL/well of serially diluted vector in the presence of 8 μg/mL polybrene. Three hours post-transduction, an additional 50 μL of fresh DMEM was added per well. Three days post transduction, medium was removed and cells treated with Tryp-LE (Sigma) for 5 min to disrupt the monolayer. Cells were then resuspended in DMEM containing a working concentration of 1/1000 of Sytox AADvanced Dead Cell Stain (Thermo Fisher Scientific). Cells were then assayed by flow cytometry using a 488 nm laser on an Attune NxT flow cytometer (Thermo Fisher Scientific). GFP expression was scored in 10,000 live cells, as determined by Sytox staining. Data were analysed with associated Attune NxT and/or FlowJo software. The % of GFP positive cells was then used to calculate vector titre of the starting material.

For CAR staining, transduced cells were trypsinised and transferred to a deep well 96-well plate. 1 mL of PBS was added per well then cells were centrifuged at 1000×*g* for 5 min. Cell pellets were then incubated with 100 μL fluorescent CAR-specific antibody (see Supplementary Table S3) diluted in 5 % Goat serum FACS stain buffer (BD) for 20 min at room temperature. Upon incubation cells were centrifuged at 1000×*g* for 5 min and supernatant discarded. Cell pellets were then resuspended in 33 % Cytofix™ buffer in PBS (BD). Fluorescence was assayed on an Attune NxT Flow Cytometer and data analysed as described above.

2.7. Vector transduction and titration by qPCR

For lentiviral vector titration by qPCR, transductions of adherent HEK293T cells were performed at 12 well plate scale, with cell densities and volumes scaled up accordingly from the 96 well plate transduction method outlined above. Transduced cells were passaged three times over 10 days before being harvested and genomic DNA extracted using an automated QIAcube DNA extraction system (Qiagen). Genomic DNA was then subjected to TaqMan™ qPCR (Thermo Fisher Scientific) using primers and probes directed to the extended packaging sequence of HIV-1 (primers 7,8/probe 1) and the cellular RPPH1 gene (primers 14,15/probe 3; see Supplementary Table S2). qPCR assays were performed on an Applied Biosystems(R) QuantStudio™ 7 thermocycler (Thermo Fisher Scientific). The average number of integrated copies per cell were used to back-calculate the integrating titres of the original vector material. For detection/quantification of cDNA formed from the MSD-to-EF1α splice acceptor junction, a custom qPCR assay (proprietary Assay IS ARWCZZ2) was developed by Life Technologies against the 'Assay 1' sequence in Supplementary Table S2.

2.8. TRiP reporter transfection

Suspension HEK293T cells were co-transfected with GFP reporter and pEF1a-TRAP expression plasmids at a molar ratio of 5:1 of reporter to TRAP. Co-transfection with pBluescript served as a no-TRAP control. Where vector genomes served as the GFP reporter, vector helper plasmids were also included in the co-transfection (see above). 48h post-transfection cells were stained with Sytox AAdvanced Dead Cell Stain (Thermo Fisher Scientific) and 10,000 live events were then assayed for GFP expression by flow cytometry using an Attune NxT flow cytometer (Thermo Fisher Scientific). For each sample GFP expression score was calculated by multiplying the median fluorescence intensity by the percentage of GFP positive cells.

2.9. Gaussia Luciferase assay

HEK293T cells were transfected in triplicate with polyA reporter plasmids using lipofection at 96 well plate scale. 48h post-transfection, media were removed from wells and transferred to a fresh plate. Remaining cell monolayers were assayed for GFP expression using an Attune NxT flow cytometer as described above. Media were assayed for luciferase activity using the *Gaussia* Luciferase Flash Assay Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Luminescence was measured using a Spectromax® i3e plate reader (Molecular Devices). For each well, *Gaussia* Luciferase activity was normalised to the corresponding GFP reporter expression in to calculate the relative *Gaussia* Luciferase activity.

2.10. RNA extraction and (q)RT-PCR

Vector production cells were washed once with PBS then total RNA was extracted using the RNeasy extraction kit (Qiagen) according to the manufacturer's instructions. RNA from virions was extracted from vector supernatants using the QiaAmp viral RNA Mini kit (Qiagen) according to the manufacturer's instructions. 500 ng of RNA was DNAse treated (EZDnase™, Thermo Fisher Scientific) according to the manufacturer's instructions. cDNA was synthesized using the SuperScript IV reverse transcription system (Thermo Fisher Scientific). Reverse transcription was performed with oligo dT or random hexamers for analysis of mRNA and U1 snRNA, respectively. cDNA was subjected to PCR using CloneAmp™ polymerase (Takara) according to the manufacturer's instructions.

For the quantitation of vector RNA splicing, RNA was DNAse-treated as above and subjected to reverse transcription using the SSIV VILO system (Thermo Fisher Scientific). cDNA was then subjected to SYBR™ Green qPCR (Thermo Fisher Scientific) using primers to detect unspliced and total vector RNA species (Primers 9, 10, 11, see Supplementary Table S2).

For 256U1[15 nt] snRNA or endogenous U1 snRNA assays RNA extracted from vector material was DNAse treated with Turbo™ DNAse (Thermo Fisher Scientific) prior to dilution. DNAse-treated RNA was then subjected to one-step qRT-PCR using the RNA-to-Ct™ system (Thermo Fisher Scientific). Primers 12,13/probe 2 used for 256U1 and primers 17,13/probe 2 used for total U1 snRNA, see Supplementary Table S2).

2.11. High-throughput sequencing

Cellular and virion RNA were extracted as per the methods outlined above. RNA sequencing was performed by GATC or Source Bioscience. Briefly, polyA-selected RNA was enriched using the NEBNext® (poly(A) mRNA Magnetic Isolation Module (NEB). The library was prepared using the NEBNext® Ultra™ II Directional RNA kit (NEB). cDNA was then subjected to 2 × 150bp Illumina sequencing to a read depth of 10M or 50M read pairs. Sequence data were then aligned to reference vector genomes using HISAT. Splice junctions were identified through analysis of split-read events with RegTools. Split reads that did not conform to canonical or semi-canonical splice events were discarded. Furthermore, introns of a size of <30 nt were also discarded from the final analysis.

2.12. Mass Spectrometry

One hundred microlitres of purified vector material was lysed and reduced in 0.1 % (w/v) RapiGest™, 50 mM ammonium bicarbonate (pH 8.0) and 5 mM dithiothreitol (DTT). Samples were incubated for 30 min at 70 °C whilst mixing. Proteins were alkylated by incubation with 15 mM Iodoacetamide (IAA) for 30 min at room temperature. Samples were then digested with Lys-C and trypsin at a ratio of enzyme:substrate of 1:10 (w/w), with incubation at 37 °C overnight. Samples were acidified using trifluoroacetic acid (TFA) to a final concentration 0.5 % (v/v) and incubated at 37 °C for 1 h on a thermomixer at 600 rpm to hydrolyse the RapiGest™ surfactant. The surfactant was removed by centrifugation at 16000×*g* for 10 min. Peptide clean-up was performed on a C18 column (Thermo Fisher Scientific) and stored in −80 °C until use. The peptides were then analysed on a U3000 nano flow HPLC (Thermo Fisher Scientific) coupled to a QExactive HF (Thermo Fisher Scientific). Tryptic peptides were injected for each sample into the LC MS/MS system operating in DIA-MS mode. DIA files were processed with the DIA-NN (V1.8) direct DIA experimental analysis workflow using default settings: carbamidomethyl (C) as fixed modification, acetylation (N-terminal) and oxidation (M) as variable modifications. Trypsin specificity was set to two missed cleavages and a protein and PSM false discovery rate (FDR) of 1 %, respectively, was chosen. Mass accuracy for precursor was set to 5 ppm, and for fragment was set to 10 ppm. Match between run (MBR) to resolve missing values and global cross run normalisation were enabled. Protein percentage composition was calculated for each protein assuming the sum of all protein intensities in each sample to be 100 %. Further statistical analyses of protein expression changes were performed in Graphpad Prism (v9.5.1) software.

2.13. T-cell transduction and killing assay

Approximately 1.5×10^6 peripheral blood mononuclear cells (PBMC) from two healthy human donors (Vial IDs: HHU20151015 [lot LP_235], HHU20170622[lot_LP327] were purchased from Cellular Technology Ltd. PBMCs were cultured with 4.5×10^6 CD3/CD28 T-Cell Expander Dynabeads™ (Thermo Fisher Scientific) in X-Vivo™ complete media (Lonza) supplemented with 10 % human AB serum (Sigma) with 100 IU/mL recombinant human IL-2 (R&D Systems). Concentrated lentiviral vectors were then added to the cells at a matched MOI. Following transduction, cells were maintained at 1.0×10^5 viable cells per mL by increasing the volume of medium. At 8 days post-transduction cells were fixed and stained for 5T4-CAR expression by flow cytometry using an AF488-AffiniPure F(ab')2 Fragment Goat Anti-Mouse IgG (Stratech) or a custom-made AF488nm-conjugated anti-CAR antibody.

T-cell killing assays were performed on THP-1 (RRID:CVCL_0006) and AML-193 (RRID:CVCL_1071) cell lines, purchased from American Type Culture Collection (ATCC; Manassas, VA, USA). CAR target antigen expression status of target cell lines was confirmed by flow cytometry using a custom-made antibody (Fleet Bioprocessing). 1×10^5 of transduced T-cells were then co-cultured in triplicate with 1×10^5 cells of each target cell line. After 24h cell culture supernatant was removed and assayed for Granzyme B release by cytometric bead array. After 40h cells were harvested and stained with CellTrace™ Far Red viability dye (Thermo Fisher Scientific). The percentage of non-viable cells were then measured using an BD™ FACSVerse™ Flow Cytometer. The presence and absence of 5T4 in THP-1 and AML-193 cell lines respectively was verified by RT-PCR (mRNA), immunoblot and flow cytometry.

2.14. SDS-PAGE and western blotting

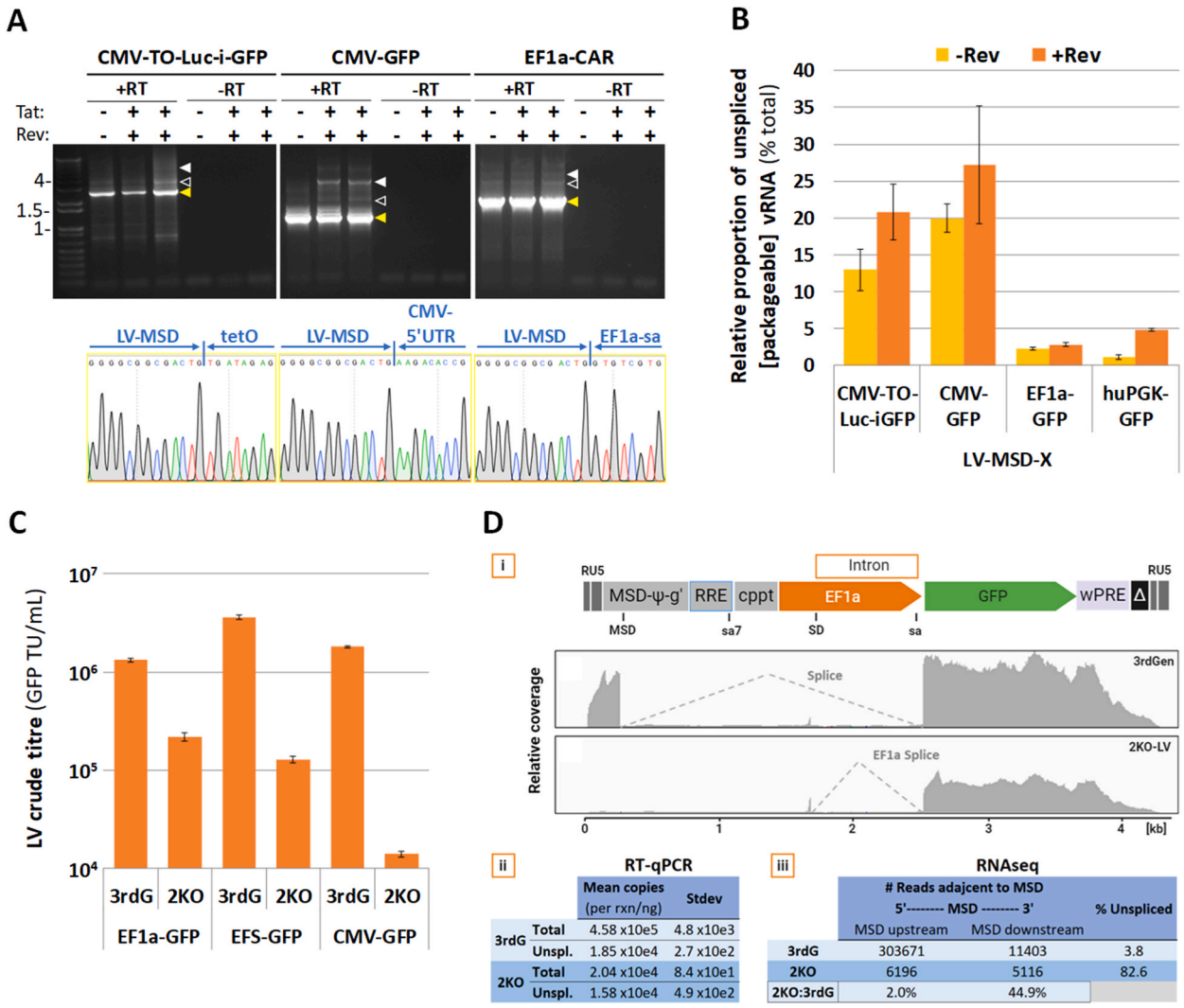
End-of-production cells were harvested in RIPA buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 1 % Triton™ X-100, 0.1 % Sodium Deoxycholate, 0.1 % SDS, 140 mM NaCl), and protein content was determined using the Pierce™ BCA protein Assay Kit (Thermo Fisher Scientific). Up to 20 μg of total protein was mixed with Laemmli Buffer (Bio-Rad), prior to resolution by SDS-PAGE. Crude vector samples were mixed with 4X Laemmli buffer with 10 % (V/V) β-mercaptoethanol to a final concentration of 1X. All Laemmli samples were incubated at 100°C for 5 min prior to loading. SDS-PAGE was performed using 4–20 % polyacrylamide gels (Bio-Rad), and samples were typically run for 2 h at 120V. Proteins were blotted onto nitrocellulose using the Trans-Blot® Turbo™ transfer system (Bio-Rad). After transfer, membranes were blocked in PBS-Tween (0.05 %) containing 5 % (w/v) milk powder for 1 h. Primary and secondary antibodies were diluted in 5 % Milk PBS-Tween. Each antibody was incubated for 1h prior to 45 min of washing in PBS-Tween. Where HRP-conjugated secondary antibodies were used, blots were incubated with SuperSignal™ West ECL reagent (Thermo Fisher Scientific) prior to imaging according to the manufacturer's instructions. Where fluorescent secondary antibodies were used, blots were visualized directly after washing. All western blots were imaged using a ChemiDoc™ MP Touch imaging system (Bio-Rad) and images were analysed using the associated software.

2.15. Other in-process residuals testing and assays

Unless otherwise noted, all assays were performed using the manufacturer's standard protocols. HEK293T host cell proteins (HCP) and HIV-1 p24 were quantified by Enzyme Linked Immunosorbent Assay (ELISA) (HCP HEK293 ELISA kit, Cygnus Technologies and Alliance HIV-1 P24 ANTIGEN ELISA Kit, Revvity Health Sciences B.V respectively), using the Dynex MXR II plate reader. The following changes were applied to the HCP assay: incubation with primary antibody was increased to 2 h with a reduced shaking speed of 200 RPM, and protein quantified against a cell lysate preparation, made in house. Total protein was assessed using BCA Micro BCA™ assay kit (Pierce Biotechnology) using a reduced total volume of 200 μL per well. Total DNA was assayed using Quant-iT™ PicoGreen™ dsDNA Reagent and Kit (Invitrogen), using the Spectramax M2e microplate reader. DNA was extracted from samples using a QIAamp 96 DNA QIAcube HT kit on the QIAcube HT system (Qiagen). Residual KanR DNA sequences were quantified by a proprietary qPCR (Taqman) assay.

2.16. Statistical analysis

Where statistical analysis is performed, data were assessed using the F-Test Two-Sample for Variances analysis to establish equal/unequal variance between data groups, followed by appropriate T-test (Student/Welch's).



		Mean copies (per rxn/ng)	Stdev
3rdG	Total	4.58 x10e5	4.8 x10e3
	Unspl.	1.85 x10e4	2.7 x10e2
2KO	Total	2.04 x10e4	8.4 x10e1
	Unspl.	1.58 x10e4	4.9 x10e2

	# Reads adjacent to MSD 5'-------- MSD -------- 3'		% Unspliced
	MSD upstream	MSD downstream	
3rdG	303671	11403	3.8
2KO	6196	5116	82.6
2KO:3rdG	2.0%	44.9%	

Fig. 2. Unwanted activity of the major splice donor leads to higher proportion of spliced RNA relative to unspliced, packageable vector genomic RNA. A) RT-PCR analysis was performed on poly-T selected RNA from post LV production cells using primers #1 and #3 (wPRE), as indicated in Fig. 1 and Supplementary Table S2. Three different transgene cassettes were assessed: CMV-TO-Luciferase-IRES-GFP (harbouring 2x tet operators), CMV-GFP and an EF1α-chimeric antigen receptor sequence (CAR). The expected sizes of unspliced, full length vRNA (white arrow) and Δψ-RRE (open arrow) are indicated. The predominant aberrant splice product is indicated by a yellow arrow; extracted DNA from these positions were sequenced, confirming the splice junction resulting from MSD splicing into transgene splice acceptors (known and cryptic). B) RT-qPCR analysis (SYBR) was performed on samples from post LV production cells using primers #9 and #10 or #11, as indicated in Fig. 1 and Supplementary Table S2. LVs encoding the stated internal Promoter-GOI cassettes were produced±pRev (data are mean [SD], n = 4). C) LV genomes were engineered to mutate the MSD and the adjacent crSD1 (see text) to produce a '2KO' LV genome (see Fig. S1 for specific sequence mutation); GFP cassettes were employed, driven by three different promoters as indicated. LVs were produced in adherent HEK293T cells (hence generally lower titres) and titrated by flow cytometry (data are mean [SD]; log10-scale, n = 2). D) Total extracted RNA from post-production cells of 3rd Gen or 2KO-LVs harbouring the EF1 α-GFP cassette was subjected to RNAseq (data are mean [SD], n = 2), [i, iii] and RT-qPCR [ii] (Some aspects created in BioRender.com).

3. Results

3.1. Generation of aberrantly spliced mRNA from LV genome cassettes is due to promiscuous MSD activity

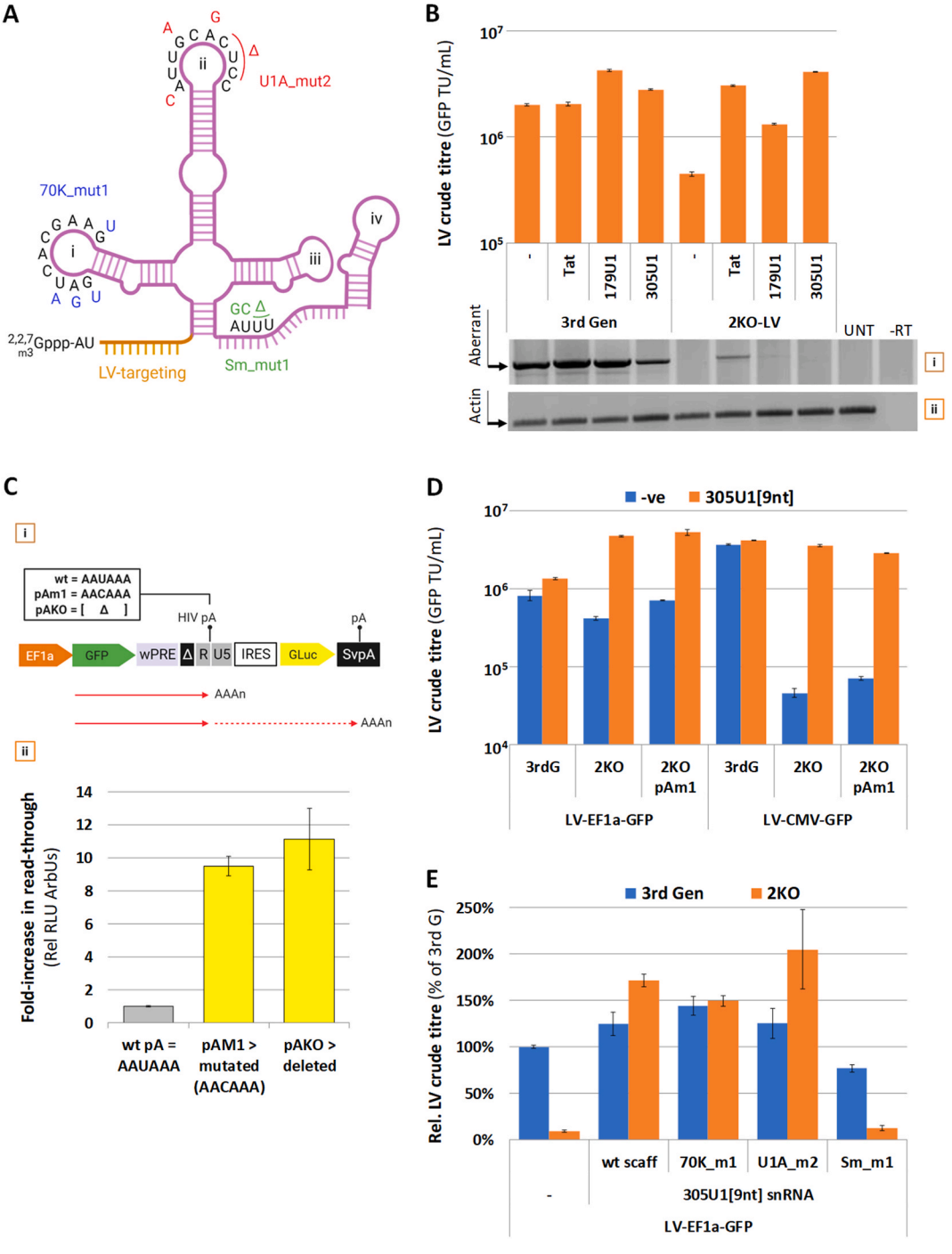
Third generation LVs based on HIV-1 incorporate the native Gag/Pol 5′UTR, encompassing RU5 and ψ sequences, through to the Gag start codon (Fig. 1). The viral MSD is located within this complex sequence, and theoretically forms an artificial, Rev-inhibited intron from the MSD through ψ, the RRE and to HIV-1 splice acceptor 7 (sa7). However, we observed that the major species of vRNA detected by RT-PCR analysis of LV production cell RNA were shorter than expected (not unspliced vRNA), and additionally did not correlate to MSD-to-sa7 spliced RNA (Δψ-RRE) (Fig. 2A). Sanger sequencing of these products revealed these species to be the result of splicing events from the MSD into the transgene 5′UTR region. This occurred regardless of the promoter or the transgene sequences used, and not only to the strong EF1α splice acceptor site but also to cryptic splice acceptors within the tet-operator or CMV 5′UTR. We developed an RT-qPCR-based splicing assay to quantify levels of total and unspliced RNA using amplicon targets shown in Fig. 1. RNA was extracted from the production cells of a panel of LV-GFP vectors, and relative levels of unspliced RNA were quantified by RT-qPCR (Fig. 2B). We observed that absolute levels of unspliced vRNA varied depending on the internal promoter-UTR sequence, and unspliced genomic RNA was a minority species of all the vectors we tested. LV genomes carrying EF1α and PGK promoters were the most severely affected, with >95 % of transcripts being spliced, although subsequently we found that for the PGK promoter, this was due to the specific 5′ UTR sequence employed. Hence most transcription events across 3rd generation LV genome expression cassettes do not lead to production of full-length vector genome RNA. In contrast, the LV-CMV-GFP vector was less severely affected, though only 25 % of vRNA remained unspliced. We found that the presence of Rev during vector production had only modest effects on the accumulation of unspliced RNA. This is supported by others' work indicating that Rev may be more important for vRNA packaging [22]. LVs employing the commonly used EF1α promoter (in therapeutic vectors) are routinely produced to high titre, and so the relatively lower abundance of full length compared to spliced vRNA appears to be sufficient to support such titres. To ablate the aberrant splicing we introduced mutations within stem loop 2 of the packaging signal, leading to inactivation of the MSD and the adjacent cryptic SD [27] (herein referred to as crSD1) to generate a '2KO-LV' genome (Supplementary Fig. S1A). Production titres of 2KO-LVs encoding EF1α-, EFS- or CMV-GFP cassettes in adherent HEK293T cells were 6-to-100-fold lower than the 3rd generation vector depending on the promoter-UTR employed (Fig. 2C), confirming the attenuating impact of MSD inactivation [20].

To provide deeper analysis, we produced 3rd generation and 2KO LVs harbouring the EF1α-GFP cassette and assayed the cellular RNA at the point of LV harvest by RT-qPCR and RNAseq (Illumina sequencing) (Fig. 2D). These data were in line with our previous results. The number of reads or copies of RNA upstream of the MSD were >20-fold greater than immediately downstream of the MSD, leading to <5 % unspliced vRNA. In contrast, the number of reads or copies of RNA upstream or downstream of the inactivated MSD within the 2KO-LV sample were approximately the same, leading to ~83 % unspliced vRNA. Intriguingly, the absolute numbers of unspliced vRNA generated by 3rd generation and 2KO LVs only varied ~2-fold by both methods using these 5' primer sets. Given the clear attenuating effect of MSD mutation on 2KO-LV output titres, this indicated some qualitative requirement for the presence of the MSD (e.g. transcription elongation), distinct from its role in (completion of) splicing.

3.2. Titres of MSD-mutated LVs are rescued by co-expression of tat or modified U1 snRNAs that target sequences in the packaging signal

There appeared to be a disconnect between the total transcriptional 'power' of the CMV promoter driving the LV genome transcription unit (reflected by the excessive amount of aberrant-spliced products), and the actual amount of full length vRNA generated when the MSD was inactive. Others have shown that the HIV-1 5′ poly-A signal (retained in LVs) is active early in infection, playing a role in premature transcription termination prior to the Tat-feed-forward loop. The recruitment of U1 snRNA to the MSD suppresses polyadenylation via telescripting at this site allowing low level transcription to occur leading to generation of Tat mRNA [25,28]. Intriguingly, a CMV-driven proviral HIV-1 genome is made Tat-dependent when the MSD is mutated, suggesting a redundant role for 5′ poly-A suppression by Tat and U1 snRNA recruitment to the MSD [29]. To investigate this possibility, we generated modified U1 snRNA expression cassettes where the splice donor complementary sequence at the 5′ end of the U1 snRNA is replaced with sequence reverse-complementary to regions upstream or downstream of the MSD, called 179U1[9 nt] and 305[9 nt], respectively (see Fig. 3A and Supplementary Table S1). The modified U1 snRNAs were transcribed by the native U1 promoter and terminator regions. Standard 3rd generation or 2KO LV encoding the EF1α-GFP cassette were produced in adherent HEK293T cells with either these LV-targeted U1 snRNAs or a Tat expression plasmid to assess the impact on LV titres and on the levels of aberrant splicing (Fig. 3B). These data show that Tat and the LV-targeted U1 snRNAs were able to increase, rescue, or even enhance 2KO-LV titres above untreated 3rd generation LV levels. Interestingly, Tat appeared to also increase the level of a background aberrant splice product, which indicated the presence of another cryptic splice site within the packaging region. The fact that the absence of Tat is a key safety aspect to 3rd generation LV platforms, and not wanting to take a retrograde step in LV development, we did not perform any further work using Tat.

Given that the data in Fig. 3B was consistent with potential roles of Tat and U1 snRNA to suppress the 5′ poly-A site within the LV cassette, we generated the '2KO-pAm1' LV genome harbouring a 5′ LTR poly-A signal mutation (AATAAA > AACAAA). Having demonstrated the efficacy of this poly-A mutation within a GFP-IRES-Luciferase reporter (Fig. 3C), we produced LV-EF1α/CMV-GFP vectors within 3rd generation, 2KO, and 2KO-pAm1 LV backbones in adherent HEK293T cells, ± p305U1[9 nt] (Fig. 3D). Strikingly, 305U1[9 nt] co-expression was able to enhance both 2KO and 2KOpAm1 LV titres by the same magnitude, and from similar low(er) baseline levels. This demonstrated that MSD mutation was not activating the 5′ poly-A, and that 305U1[9 nt] snRNA was not acting to suppress such activity. Moreover, the testing of mutant 305U1[9 nt] snRNAs bearing changes in binding sites for known protein-U1 interactors U1-70K (involved in poly-A suppression [30]) and U1A (Fig. 3A), showed that such functions are not necessary for the



(caption on next page)

Fig. 3. Tat or LV-targeted U1 snRNAs co-expressed during production of 3rd generation or 2KO-LVs increase output titres by a mechanism independent of the 5′ poly-A site. A) The cloverleaf structure of U1 snRNA (stem loops i to iv) and the complementary LV-targeting sequence in place of the native splice donor annealing sequence. Nucleotide changes for mutant LV-targeted U1 snRNAs for 70K, U1A and Sm protein binding site ablation are indicated. B) Output titres for 3rd generation or 2KO-LVs (EF1α -GFP) generated by co-transfection of adherent HEK293T cells with pTat, p179U1[9 nt] or p305U1[9 nt] are displayed (data are mean [SD]; log10-scale, n = 3), alongside RT-PCR products from RNA extracted from associated post-production cells; [i] aberrant MSD-EF1a splice acceptor product and [ii] actin mRNA (control) product is indicated. RT-PCR carried out using primers #1 and #4, as indicated in Fig. 1 and Supplementary Table S2. C) Validation of native poly-A mutations within the R-U5 sequence by testing of a GFP-IRES-Luciferase(*Gaussia*) reporter construct in HEK293T cells. The (SIN)R-U5 sequence harbouring either intact or a point mutant or deleted polyA signal was inserted between GFP and GLuc ORFs [i], and therefore the relative transcriptional read-through was measured by GFP-normalised luciferase activity [ii] (SvpA; SV40 poly-A) (data are mean [SD], n = 3). D) Output titres of 3rd generation LVs and 2KO-LVs (with indicated GFP cassettes) bearing either the intact or pAm1-mutated 5′ poly-A site, produced from adherent HEK293T cells co-transfected ± p305U1[9 nt] (data are mean [SD]; log10-scale, n = 2). E) Output titres of 3rd generation LVs and 2KO-LVs containing the EF1α-GFP cassette produced from adherent HEK293T cells co-transfected ± LV-targeted 305U1[9 nt] snRNAs with either wild type (wt) scaffold or the stated loop mutants for 70K, U1A or Sm protein binding denoted in A) (data are mean [SD], n = 2). (Some aspects created in BioRender.com).

increase in titres (Fig. 3E). The only requirement was for the Sm-protein binding site, suggesting that correct processing of LV-targeted U1 snRNAs is required. The data also highlighted the interesting differences between 2KO-LV genomes harbouring different promoters (also seen in Fig. 2C); namely, that the attenuating impact of MSD inactivation was less within the EF1α-driven cassette, although perhaps more variable. Also of note, was the surprising modest effect on GFP expression of the lack of a transgene poly-A site within the 2KO-pAm1 vector, given that the pAm1 mutation is copied to the 3′LTR after cDNA synthesis. We only observed a modest reduction in GFP expression of cells transduced by the 2KO-pAm1 vector, which still enabled titration by flow cytometry. Thus, transcription termination of the GFP pre-mRNA in this case must occur by use of cryptic or genuine poly-A site(s) within the integrated cellular gene transcription unit, leading to longer 3′ UTRs for the transgene mRNA. This finding has led us to develop more potent poly-A sequences within SIN-LTR to improve transcriptional insulation and transgene expression in target cells (reference [31] and manuscript under review).

3.3. Maximal enhancement of 2KO-LV output titres by optimisation of LV-targeted U1 snRNAs and MSD-inactivating sequences

Next, we sought to develop the approach of utilising LV-targeted U1 snRNAs to enhance 2KO-LV titres. We tested the effect of increasing the transfected amount of p179U1[9 nt] or p305U1[9 nt] from 10 to 2000 ng/mL (Supplementary Fig. S2B). These data indicated that the specific target site rather than being able to maximise modified U1 snRNA levels was the most important parameter; 305U1[9 nt] being able to mediate a rescue in titres at the lowest input level, in contrast to 179U1[9 nt], which did not enable full rescue at any input amount level. Next, we generated additional 305U1 snRNA variants with increasing length of complementarity to the target RNA (see Supplementary Table S1), and assessed their effect on 2KO-LV-EF1α-GFP titres (Supplementary Fig. S2A). These data indicated that a target length of minimally 10 or longer nucleotides enabled the greatest rescue of 2KO-LV titres. Using this knowledge, we designed a panel of re-targeted U1 snRNAs against the full vector packaging/backbone sequence, with all variants having 15 nucleotides of target complementarity, and assessed their ability to enhance 2KO-LV-EF1α-GFP titres produced in adherent HEK293T cells (Supplementary Fig. S2C). Targeting the core packaging sequence from SL1 to SL3 provided the most pronounced increase in 2KO-LV-EF1α-GFP titre.

To develop the 2KO-LV genome towards gene therapy product manufacturing, we transitioned our studies from adherent cells into suspension (serum-free) HEK293T production cells. We repeated the LV-targeted U1 snRNA panel screen in HEK293T suspension cells at small scale during production of 3rd generation or 2KO LVs encoding EF1α-GFP (Fig. 4A) or EFS-GFP cassettes (Supplementary Fig. S2D). This confirmed that re-targeting modified U1 snRNAs to the core packaging sequence produced the greatest recovery/boost to 2KO-LV output titres. Variant 256U1 generally provided the greatest effect irrespective of the type of MSD-inactivation (‘2KO’ is a subtle mutation, whereas ‘2KOm5’ is a full SL2 loop replacement - see later) or the internal promoter employed.

Next, we returned to the observation that specific 2KO mutation did not entirely ablate aberrant splicing from the core packaging region, as evidenced by the increase in spliced product in the presence of Tat (Fig. 3B). This result is consistent with the activation of cryptic splice donor sites within Ψ when inactivating the MSD. This is also the case when the adjacent crSD1 site is mutated in the 2KO variant since we identified two further sites (crSD2 and crSD3) within the SL4 loop (see Supplementary Fig. S1A). We generated a series of mutants within the SL2 and SL4 regions of the core packaging region (see Supplementary Fig. S1A), where the MSD was mutated by default (all variants 1KO-to-4KO) and additionally, further variants harboured differing mutated cryptic splice donor sites. This produced mutants ‘2KO’ (crSD1-KO; the initial variant tested), ‘3KO’ (crSD1,2-KOs) and ‘4KO’ (crSD1,2,3-KOs). We also generated a series of variants that combined SL4 crSD mutations with an entire SL2 replacement called ‘2KOm5’ displayed in Fig. 4B. The 2KOm5 modification was engineered to maximise complementarity to endogenous U1 snRNA (whilst not encoding a *bona fide* splice donor site), and additionally such that secondary structure in the core packaging signal was expected to be retained, as overlayed onto two published models (Fig. 4B i-iii) [32,33]. The native SL2/MSD sequence harbours 20 hydrogen bonding (H-bond) interactions with endogenous U1 snRNA. Under the premise that some ability to recruit endogenous U1 snRNA to the 2KOm5 sequence may be of benefit to MSD-mutated LVs, the 2KOm5 variant is able to equally match these number of H-bonds. The initial 2KO variant has only 10 predicted H-bonds with U1 snRNA (Supplementary Fig. S1B). Production of these variants within LV-EF1α-GFP vectors alongside the 3rd generation LV vector was performed in suspension HEK293T cells, co-expressing ± 256U1[15 nt] snRNA. This confirmed both the

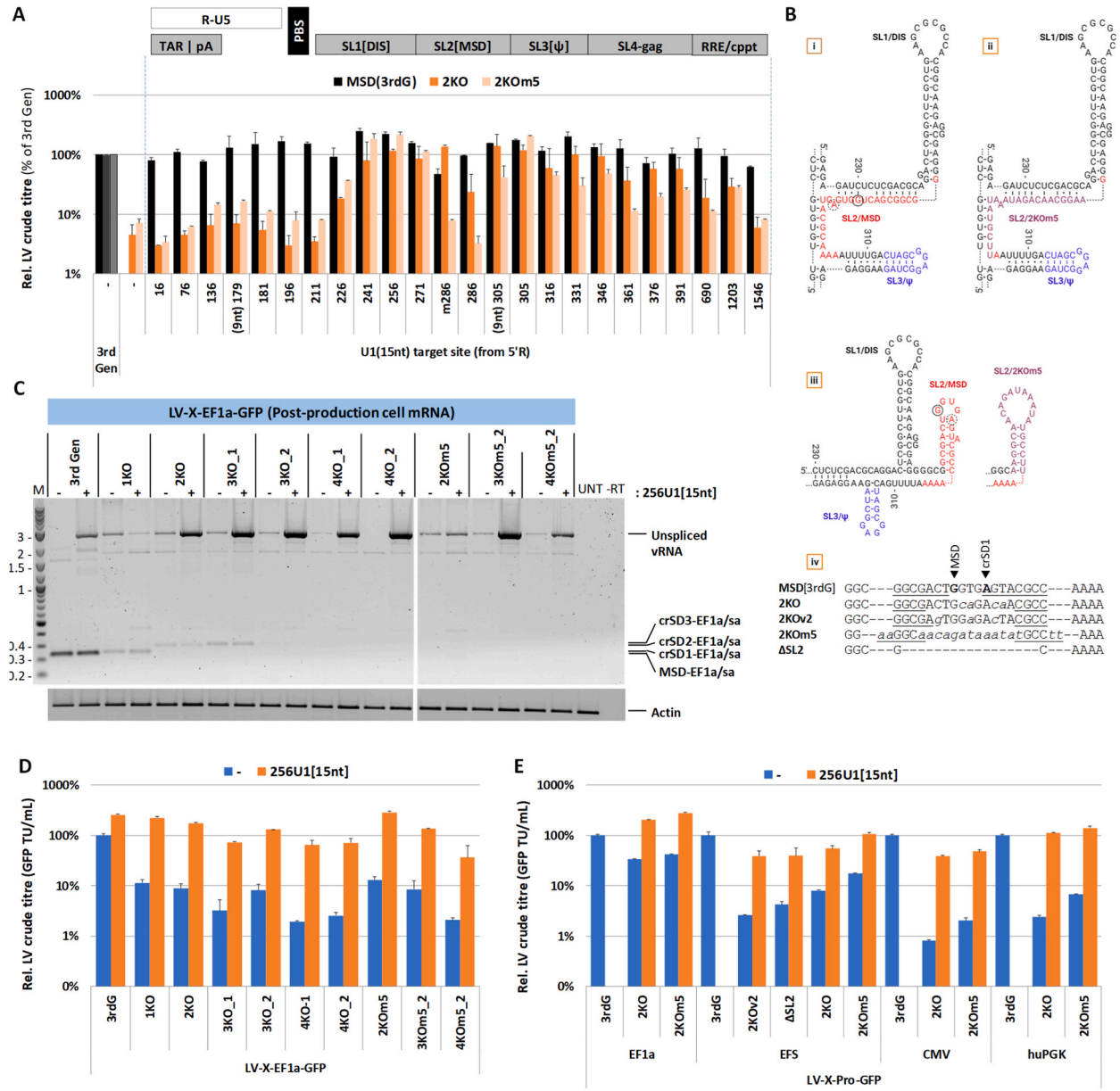


Fig. 4. Optimisation of LV-targeted U1 snRNAs and modifications to the SL2-ψ region of 2KO LV genome. A) The results of a panel screen of modified U1 snRNAs targeted to 15 nucleotide sequences across the LV packaging region (see Table S1). The approximate target region of each modified U1 snRNA is displayed schematically above the graph, indicating the position of the R-U5 sequence (with TAR and poly-A loops), primer binding site (PBS), and SL1-4 loops of the core packaging signal, as well as the RRE/cppt region (not to scale). LV-targeted U1 snRNAs were co-expressed during production of 3rd generation or MSD-inactivated ('2KO' or '2KOm5') LV-EF1α-GFP in suspension HEK293T cells, and LVs titrated by flow-cytometry. Titres are plotted relative to the equivalent 3rd generation LV produced alone (data are mean [SD], log10-scale, n = 2). B) Panels [i] – [iii]; predicted RNA structure of the core packaging signal within the 2KOm5 variant compared to 3rd generation LV by overlaying sequences onto published structures (see text for references). MSD and crSD1 sites are circled in solid and dotted lines respectively. Panel [iv] describes variant 2KO mutants relative to the unmodified SL2 sequence. C) RT-PCR analysis of extracted mRNA from post-production cells of LVs with 3rd generation or MSD-inactivated LV variant-encoding LV genomes made±256U1[15 nt]. RT-PCR carried out using primers #1 and #4, as indicated in Fig. 1 and Supplementary Table S2. Specific MSD-inactivating mutations/modifications are described in Supplementary Fig. S1. Spliced products are produced from the MSD or three activated cryptic splice donors to the EF1α splice acceptor. D) Relative GFP titres of LVs generated in C), with the 3rd generation LV titre set to 100 % (data are mean [SD], log10-scale, n = 2). E) Relative GFP titres of LVs harbouring specific MSD-inactivating mutations outlined in B) panel [iv], with the relevant 3rd generation LV titre set to 100 % (data are mean [SD], log10-scale, n = 2). (Some aspects created in BioRender.com).

substantial levels of aberrant splicing from MSD to the EF1α splice acceptor in the 3rd generation construct, but that inactivation of the three other crSDs (1,2,3) was necessary to fully suppress aberrant splicing from this region (Fig. 4C). Interestingly, the 2KOm5 variant appeared to suppress aberrant splicing from the entire region, despite not containing mutations in crSD2 or crSD3. The 2KOm5 LV was the least attenuated in terms of output titres, and was maximally increased to levels greater than the 3rd generation LV when 256U1

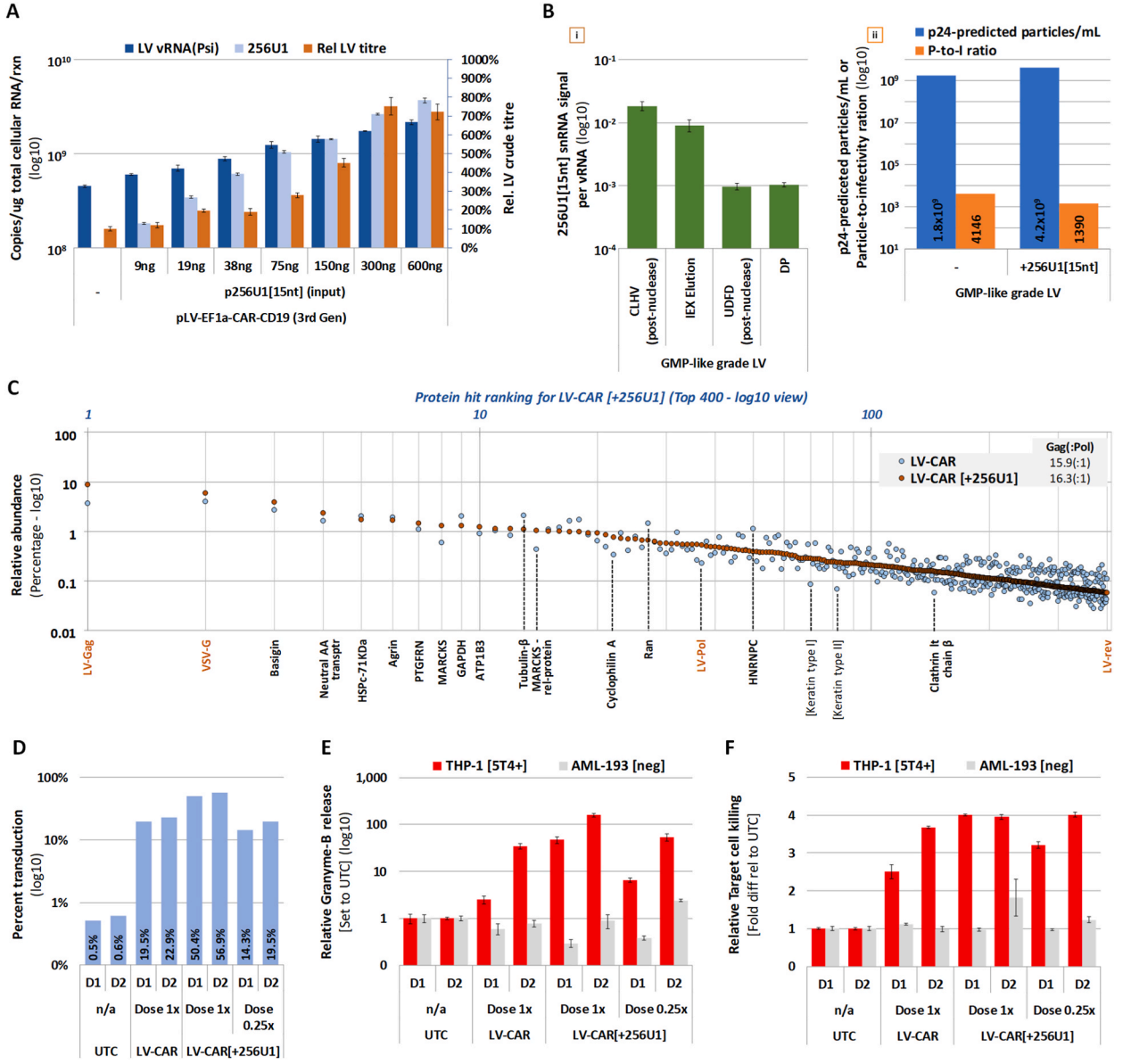


Fig. 5. The use of 256U1 LV enhancer to in production of 3rd generation LVs for CAR-T therapy. A) The dose response effect of titrating-in p256U1 plasmid during co-transfections of suspension HEK293T cells resulting in enhanced LV-CAR-CD19 output titres (HEK293T integration titres; relative to no enhancer [100 %], data are mean [SD], n = 4). Extracted total RNA was analysed for vRNA (ψ; primers #7, #8, probe #1) and 256U1 snRNA (primers #12, #13, probe #2) copies/reaction via RT-qPCR (data are mean [SD], log10-scale, n = 2). B) Panel [i]; detection of residual 256U1 snRNA signal (RT-qPCR) through Downstream processing of LV-EF1α-CAR-5T4 in a GMP-like process; 256U1 snRNA signal (copies) plotted as a ratio to vRNA (ψ) using same primers/probes as A) (data are mean [SD], log10-scale, n = 8). Panel [ii] the final drug products (DP) from [i] were analysed for physical particle analysis (p24-predcited particles/mL) to yield particle-to-infectivity (P-to-I) ratios. C) Mass Spectrometry protein analysis of R&D-tech grade LV-CAR-5T4 produced ± 256U1 snRNA. The top 400 protein hits are plotted (rank position Log10 scale) against relative abundance (% of total; Log10 scale), and key proteins highlighted. Data representative from triplicate runs. D) Activated PBMCs from two donors (purchased from Cellular Technology) were transduced with stated relative doses of LV-CAR-5T4 ± 256U1 snRNA and percent CAR expression determined by flow cytometry. E) Untransduced T-cells (UTC) or CAR-T cells from D) were incubated with 5T4-positive or 5T4-negative cell lines and release of Granzyme-B measured (plotted relative to UTC, set to 1. Data are mean [SD], log10-scale, n = 3). F) Killing of 5T4-positve or 5T4-negative cell lines from D) was measured by cell viability counts and plotted as fold-killing relative to UTC. Data are mean [SD], n = 3.

[15 nt] snRNA was co-expressed (Fig. 4D). This intimated that the ability to recruit endogenous U1 snRNA to the SL2 region lessened the effect of MSD mutation. The effect of co-expression of 256U1[15 nt] on the amount of RT-PCR product pertaining to unspliced vRNA was generally correlated to LV output titre, with the caveat that RT-PCR is semi-quantitative as used here. Clearly the (apparent) maximal cellular unspliced vRNA expression did not lead to maximal LV titre in all cases, and so we couldn't exclude the possibility that some modifications to the Ψ region impacted the packaging potential of some of these variants. Nevertheless, since the 2KOm5 variant produced the greatest titre (plus 256U1[15 nt]) without generating aberrant splice products, we moved forward with this variant. We then confirmed these results in the context of other internal promoters, with the 2KOm5 variant performing the best compared to three other MSD-mutations (including an SL2-deletion; Fig. 4B panel iv), and the provision of 256U1[15 nt] snRNA restored LV output titres to similar levels as the 3rd generation LV (Fig. 4E).

3.4. LV-targeted U1 snRNAs can enhance the production and particle-to-infectivity ratio of 3rd generation LVs

The use of certain LV-targeted U1 snRNAs appeared to increase the output titres of 3rd generation LVs (i.e. with an intact MSD), particularly those containing the EF1α internal promoter (See Fig. 3B, D, and 4B, 4D). To develop Ψ-targeted U1 snRNAs as general enhancers of LV production, we generated an RT-qPCR assay to detect 256U1[15 nt] snRNA within process samples such that the signal could be unambiguously distinguished from endogenous U1 snRNA (Supplementary Fig. S3A and 3B). We measured the amount of 256U1[15 nt] snRNA, vRNA (primer/probe set within Ψ, downstream of MSD) and LV output titres during production of a 3rd generation LV encoding EF1α-CAR19 (Chimeric antigen receptor targeted to CD19) in suspension HEK293T cells, where the amount of p256U1[15 nt] was progressively increased. A clear positive correlation between 256U1[15 nt] snRNA, vRNA and output titres was observed; a plasmid input of between 150 ng and 300 ng (12.5–25 % total pDNA) produced the maximal effect (Fig. 5A).

Next, we tested whether 256U1[15 nt] snRNA remained associated with packaged vRNA, although this seemed unlikely since 256U1 snRNA targets across the ψ dimerization signal. We produced both 'tech' grade (Supplementary Fig. S4A) and 'Good Manufacturing Practice (GMP)-like' grade (Fig. 5B[i]) LV-CAR5T4 vector material (encoding a CAR targeting the protein 5T4). We tested the relative clearance of 256U1[15 nt] snRNA signal through each concentration/purification process. These results show that 256U1[15 nt] snRNA signal within both LV material types behaves similarly to host cell nucleic acid clearance through processing, being sensitive to nuclease treatment. The tech grade LV stock, in which the final concentration step involved centrifugation and pellet-resuspension, had an vRNA-to-256U1[15 nt] ratio of 32:1. The GMP-like LV stock, which included ultra-diafiltration after the second nuclease treatment, had an vRNA-to-256U1[15 nt] ratio of ~1000:1. In a repeated production, this time of an LV-CARCD19 vector with 256U1[15 nt], we compared the clearance of other process DNA and protein residuals through a commercially relevant process, including 'total' U1 snRNA (i.e. endogenous plus 256U1[15 nt] snRNA), relative to LV components (Supplementary Fig. S4C–H). These data demonstrated that residual 256U1[15 nt] signal is cleared more comparably with residual DNA and endogenous U1 snRNA than other residuals.

We further quantified LV-CAR5T4 particle production by the GMP-like process (± 256U1[15 nt]) and observed that both particle production and the particle-to-infectivity ratio of LV produced plus 256U1[15 nt] was improved by 2-to-3-fold (Fig. 5B[ii]), resulting in a net increase of 7-fold output titre (data not shown). We performed Mass Spectrometry analysis on the 'tech' grade LV material and compared the relative differences in abundance of the top 400 proteins detected (Fig. 5C). This analysis revealed the following [1]: the expected abundances of LV protein components relative to each other (a Gag-to-[Gag]Pol ratio of ~16:1; for HIV-1, ratio is 10–20:1 [34]) [2]; the presence of known host cell factors that are co-packaged (e.g. Cyclophilin-A); and [3] that the expression of 256U1[15 nt] snRNA during production had not resulted in incorporation of any unwanted/unexpected cellular protein(s) due to potential off-target effects of the modified U1 snRNA on gene mRNA production.

We used the 'tech' grade LV-CAR5T4 stocks to transduce peripheral blood mononuclear cells (PBMCs) from two healthy donors to assess any potential impact of 256U1 snRNA residual signal on the *in vitro* potency of resultant CAR-T cells at day 8 post-transduction. Both LV stocks transduced PBMCs to similar levels (Fig. 5D), and resultant CAR-T cells secreted Granzyme-B well above background levels (Fig. 5E), causing similar levels of 5T4-positive cell killing (Fig. 5F). We included a parallel transduction with LV-CAR5T4 produced plus 256U1[15 nt] at 25 % of the primary dose, to provide additional data regarding residual 256U1[15 nt] snRNA signal clearance in CAR-T cell cultures (Supplementary Fig. S4B). This allowed comparison of stability of residual vRNA (protected within LV capsids) and 256U1[15 nt] snRNA in culture over a seven-day period. In comparison to the cellular mRNA transcript for *RPPH1*, the relative rate of decline of 256U1[15 nt] snRNA signal was ~4.5-fold greater than that of vRNA, indicating that the 256U1 [15 nt] snRNA was likely more exposed to extracellular nucleases. This was consistent with the previous data showing that this signal mostly derived from production residuals rather than being embedded within the LV particles. Whilst the 0.25x dose of LV-CAR5T4 produced plus 256U1[15 nt] achieved similar levels of PBMC transduction, we couldn't exclude the possibility of a saturation effect at the 1x dose for both LV preps. However, due to the ~7-fold difference in integrating titres of the two LV stocks (due to the enhancement by 256U1[15 nt]), this allowed 7-fold lower volumetric dosing with LV made plus 256U1[15 nt] snRNA. Therefore, PBMCs would have been exposed to 7-fold less residuals, which may be beneficial to post-transduction recoveries in CAR-T cells.

Next, we assessed the utility of these LV-targeted U1 snRNAs to enhance the production of 3rd generation LVs harbouring more complex transgene cassettes. First, we generated a typical β-Globin LV genome used in the treatment of Sickle cell disease and β-Thalassaemia [35], containing an inverted β-Globin gene cassette (Supplementary Fig. S5A). Production titres of this β-Globin LV could be enhanced by 2-to-3-fold by co-expression of LV-targeted U1 snRNAs (Supplementary Fig. S5B). Further, we generated an LV genome encoding alpha-1 anti-trypsin (αAT; used for the treatment of Alpha-1 Antitrypsin Deficiency; AATD [36]) fused to GFP via a T2A self-cleaving peptide (*Thoseaasigna* virus), with codon-optimised or native ORFs (Supplementary Fig. S5A). LVs were pseudotyped with either VSV-G or the trypsin-dependent Sendai virus F/HN. Co-expressed 256U[15 nt] snRNA increased VSV-G pseudotyped LV

titres by 2-to-3-fold, and surprisingly, a 17-fold increase in production of the Sendai virus F/HN pseudotyped LVs was observed (Supplementary Fig. S5C). Whilst the absolute LV-SeV-F/HN titres were generally low (LV-CMV-GFP_SeV-F/HN titres were 30-fold lower than LV-CMV-GFP_VSV-G), we sought to uncover a potential mechanism for the relatively larger effect of 256U1[15 nt] snRNA on SeV-F/HN LVs. An immunoblot of post-production cells against the transgene proteins revealed a modest but consistent reduction in transgene protein expression when 256U1[15 nt] snRNA was co-expressed. Given the necessary step of trypsin-activation of SeV-F protein, these results are consistent with αAT-dependent inhibition of trypsin-activation of the SeV-F/HN envelope on the LV particles produced. Therefore, the lower expression of secreted αAT produced in the presence of 256U1[15 nt] likely led to a more efficient envelope-activation step. The mechanism of how 256U1[15 nt] inhibits transgene expression is discussed later.

3.5. *Spliced vRNAs within 3rd generation LVs are templates for production of episomal cDNAs in target cells*

We assessed the consequences of the presence of spliced vRNAs within 3rd generation LV product. We produced 3rd generation LV-EF1α-GFP and 2KO[m5]-LV-EF1α-GFP vectors and treated crude harvest with Benzonase™ before purifying the LVs by ion-exchange chromatography followed by a treatment with salt-active nuclease (SAN). We confirmed the presence of MSD-EF1α spliced vRNA only within the 3rd Gen LV product by RT-PCR of extracted mRNA from production cells, as well as the crude and concentrated LV product (Fig. 6A). The 'no RT' controls carried out on the concentrated/purified LV stocks demonstrated the absence of contaminating DNA. Next we transduced adherent HEK293T cells with the concentrated/purified stocks of the two vectors at three MOI-matched doses of 1x, 2.5x and 5x. Transduced cell cultures were passaged for 10 days, taking total genomic DNA samples throughout for qPCR against Psi (ψ-vRNA; 5′ target), wPRE (3′ target) and the MSD-EF1α splice acceptor junction (see Fig. 1) at days 3, 6 and 10. By day 10, both vector stocks achieved vector-copies per cell of ~2, ~5 and ~10 for the three doses respectively, demonstrating that functionally equivalent doses of LV stocks had been applied in each case (Fig. 6B). These data show that cDNA resulting from the MSD-EF1α spliced vRNA product was readily detected in the HEK293Ts cells transduced with the 3rd generation LV vector at day 3 (Fig. 6B). However, this signal rapidly diminished during passaging and became almost undetectable by day 10 post-transduction, indicating that this cDNA was episomal. Interestingly, the total wPRE copy-number (reflecting both full length and spliced vRNA/cDNA) for the 3rd generation LV-transduced cells correlated with the sum of Psi and MSD-EF1α spliced copy-numbers at day 3 post-transduction, and then aligned with Psi copy-number once the MSD-EF1α spliced cDNA had diminished at days 6–10 post-transduction. This suggests that the MSD-EF1α spliced cDNA accounted for most, if not all episomal cDNA generated immediately after transduction. In contrast, the copy-number of Psi and wPRE cDNA for the 2KO-LV transduced cultures were equivalent within each dose, and remained stable throughout the time course, with no MSD-EF1α spliced cDNA being detected.

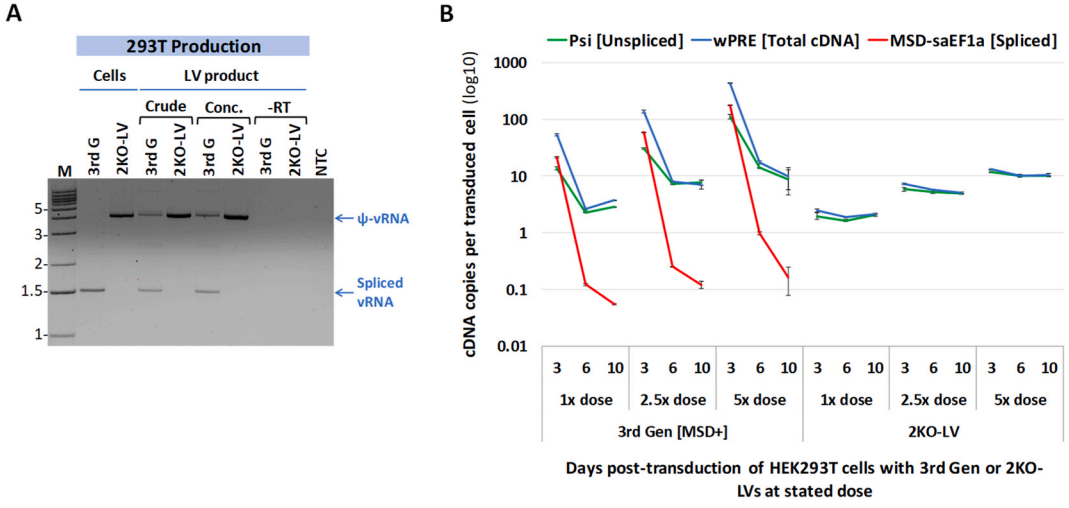


Fig. 6. Spliced vRNA is detect in 3rd generation LV product and is converted to episomal cDNA in target cells. LV-EF1α-GFP vectors produced in 3rd generation or 2KO[m5]-LV (+256U1[15 nt]) genomes by transfection of suspension HEK293T cells followed by tech scale purification/concentration and transduction of adherent HEK293T cells. A) RT-PCR analysis of extracted mRNA from post-production cells ('Cells') or LV product (crude harvest or concentrated). Full length, packageable LV vRNA (ψ-vRNA) and spliced vRNA (MSD-to-EF1α splice acceptor) are denoted; RT-negative (-RT) was performed on concentrated LV product only. RT-PCR carried out using primers #1 and #2, as indicated in Fig. 1 and Supplementary Table S2. B) Concentrated LV stocks were used to transduce adherent HEK293T cells at matched doses achieving similar integrated copies of both vector types by day 10 post-transduction. Cultures were passaged three times over a 10 day period, taking total DNA for qPCR analysis for Psi (primers #7, #8, probe #1; Unspliced – 5′ of MSD), wPRE (primers #16, #17, probe #4; Total – 3′ of MSD) and MSD-saEF1a (Assay 1; Spliced junction). Copy numbers were normalised to the RRPH1 cell control (primers #14, #15, probe #3); data are mean [SD] (technical duplicates); log10-scale; representative result from two independent experiments.

3.6. MSD-mutation and an improved TRiP System™ lead to maximal transgene repression within HIV-1 based LVs

The TRiP System™ is a vector agnostic approach to minimising unwanted effects of the therapeutic protein in production, processing, and product, and has been implemented across Lentiviral, AAV and Adenoviral systems [23]. It utilises the bacterial protein

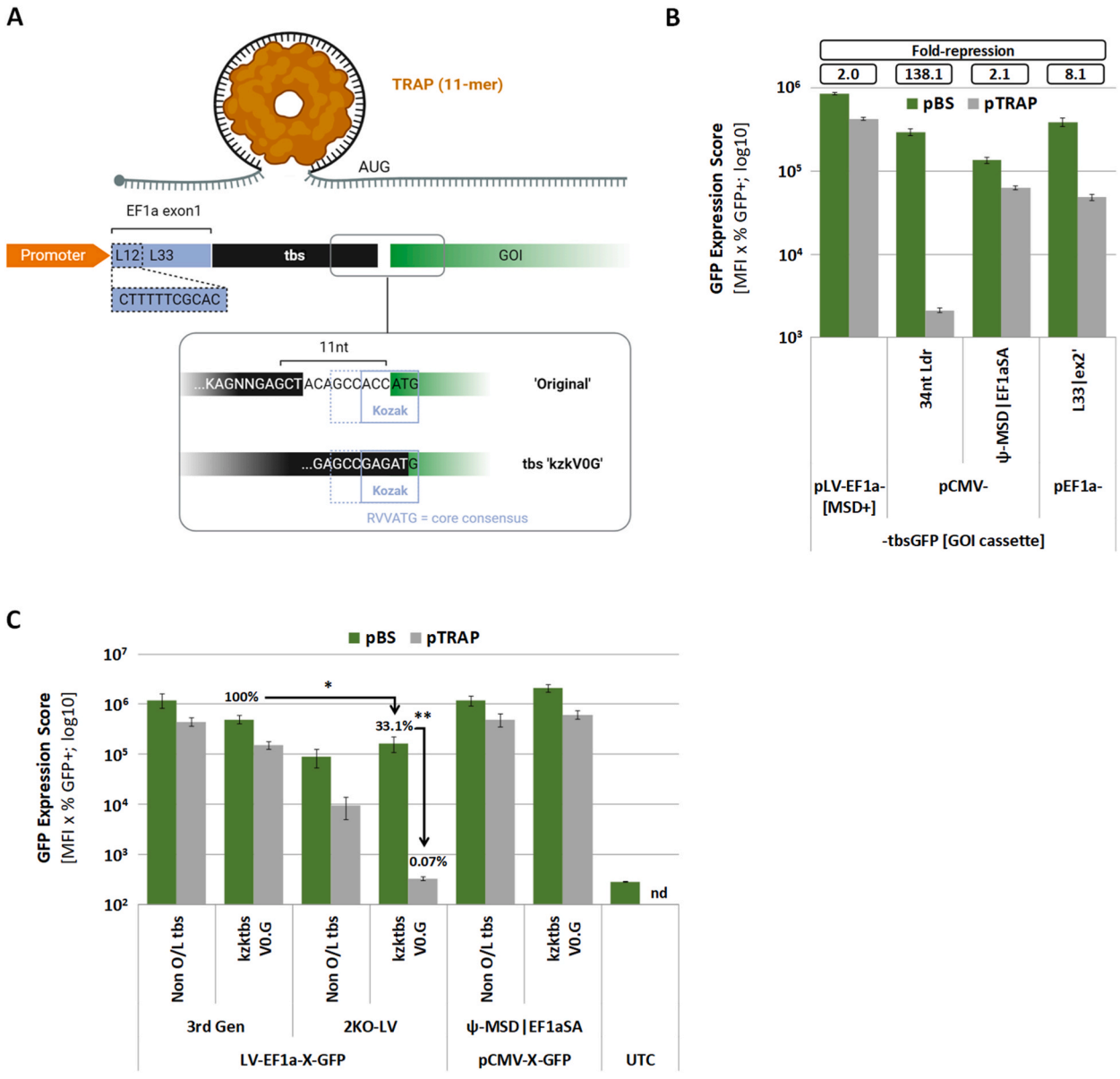


Fig. 7. Implementation of an improved TRiP 5′UTR-tbs sequence within 2KO-LVs enables greater transgene repression. A) The TRiP system employs the bacterial protein TRAP to repress translation initiation of the transgene by binding to its RNA binding site (tbs) of 55 nucleotides (consensus [KAGNN]x11) around the surface of the toroidal TRAP 11-mer. The optimal leader L12 is shown (L33 = EF1α exon 1), as well as the configuration of the kzktbsV0 variant, where the 3′ end of the tbs overlaps with the Kozak sequence. B) Comparison of gene expression of tbs-GFP cassettes (1st generation TRiP configuration) within reporter plasmids and a 3rd generation LV-EF1α-plasmid in co-transfection of suspension HEK293T cells with pBlueScript or pTRAP (GFP Expression Scores - ArbUs; data are mean [SD]; log10-scale data, n = 3). The predicted MSD-EF1α-splice acceptor spliced mRNA product was cloned into a pCMV reporter plasmid. C) Testing of the kzktbs-V0.G variant compared to the non-overlapping tbs in 3rd generation and 2KO-LV genomes (EF1α-GFP), and in the pCMV-GFP reporter with the MSD-EF1α spliced leader. Suspension HEK293T cells were co-transfected with stated constructs with pBlueScript or pTRAP and GFP Expression scores generated (ArbUs). The GFP ES score drops from 100 % to ~33 % when moving from 3rd generation LV to 2KO-LV (GFP mRNA with MSD-EF1α spliced leader no longer generated), and down to 0.1 % in the presence of TRAP. A reporter generating a GFP mRNA with MSD-EF1α spliced leader is poorly repressed by TRAP irrespective of Kozak/tbs configuration (nd – not done). *p = 0.00039 (Student T-Test, equal variance), **p = 0.0044 (Welche's T-Test, unequal variance), n = 4. Data are mean [SD]; log10-scale (Some aspects created in BioRender.com).

tryptophan RNA-binding attenuation protein (TRAP) to block translation initiation when bound to its 55 nucleotide 'TRAP binding sequence' (tbs) present in the 5′UTR, adjacent to the downstream Kozak sequence (Fig. 7A). During our wider implementation of the TRiP System™ to HIV-1 based LVs, we found that repression levels varied according to internal promoter-5′UTR sequences, and indeed these studies revealed that MSD-derived spliced vRNA products partly contributed to this variability. To confirm this, we generated a GFP reporter plasmid containing the predicted mRNA product generated by the aberrant MSD-to-EF1α splice acceptor splicing event from a 3rd generation LV-EF1α-tbs-GFP genome, and compared the repression levels by TRAP alongside our first generation TRiP GFP reporter plasmid with 'original' 5′ UTR-tbs configuration (pCMV-34ntLdr-tbsGFP; Fig. 7B). This confirmed that TRAP was minimally able to repress translation of the spliced vRNA product, indicating that sequences upstream of the tbs (in this case part of ψ) can impact TRiP efficacy. This led us to make further improvements to the TRiP System™ that are beyond the scope of this work (reference [37] and manuscript in preparation). In summary however, we generated a 'universal' 5′UTR-tbs leader – referred to as 'kzktbsV0' (or tbsV0) - composed of a twelve-nucleotide leader upstream of the tbs, and designed the 3′ end of the tbs to overlap the Kozak sequence in order to better occlude the primary AUG codon from the translation initiation machinery (see Fig. 7A). We compared the original 5′ UTR-tbs leader and the kzktbsV0.G variant for GFP repression within a 3rd generation or 2KO-LV genome (EF1α-tbs-GFP cassette), and

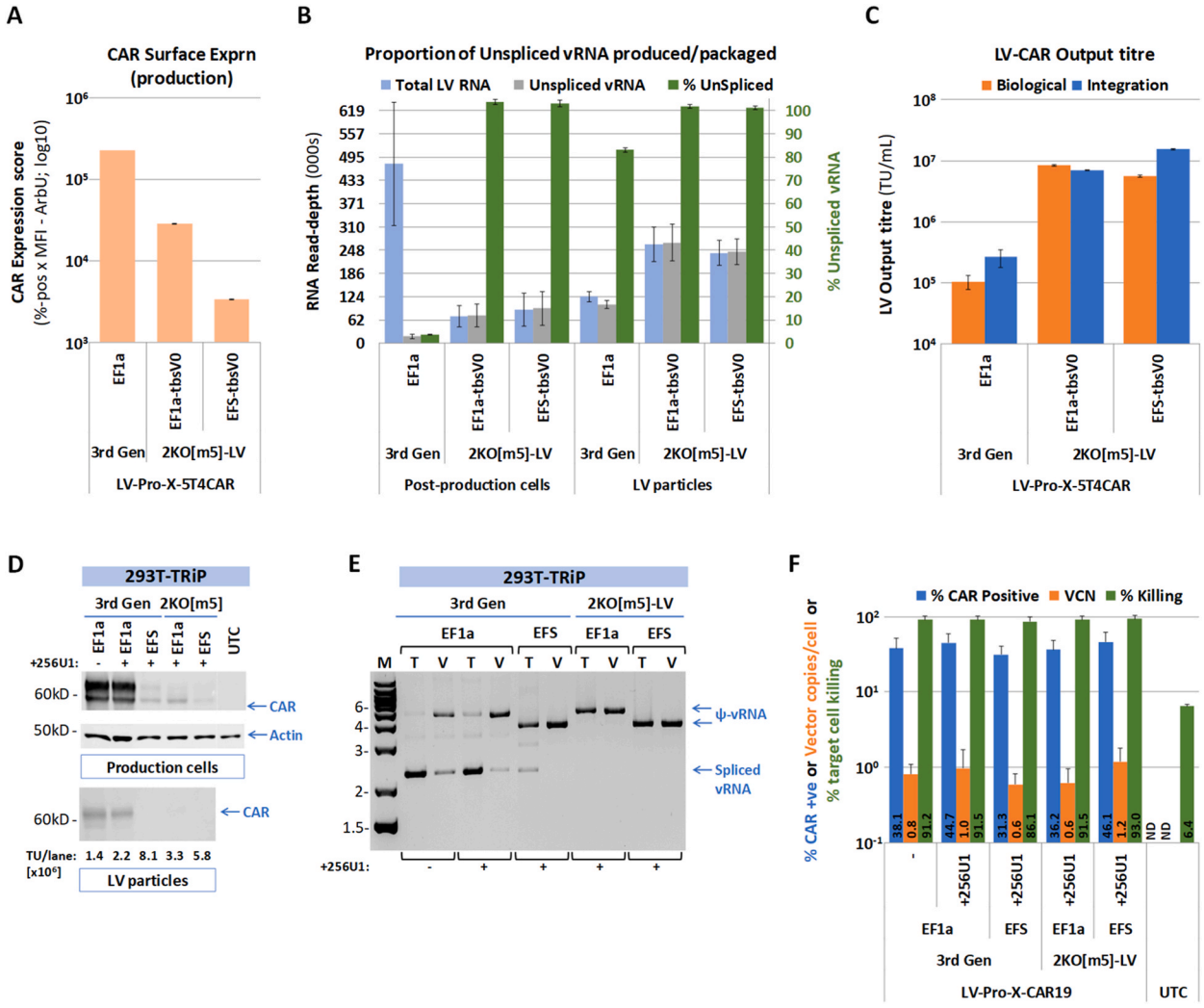


Fig. 8. Comparisons of 3rd generation LV and 2KO-LV/TRiP vector genomes in LV-CAR production. CAR-5T4 encoded 3rd generation or 2KO-LV vectors were produced in suspension HEK293T-TRiP cells, where the following analyses were carried out: A) surface CAR Expression scores (%-positive x MFI; data are mean [SD]; log10-scale data, n = 2), B) vRNA splicing analysis of total or cytoplasmic cell RNA, and vector RNA by RNAseq; data are mean [SD]; log10-scale, n = 3. C) output titres of LVs by CAR/flow cytometry (Biological) and integration assay in adherent HEK293T cells (data are mean [SD]; log10-scale, n = 2). CAR-CD19 encoded 3rd generation or 2KO-LV vectors were produced in suspension HEK293T-TRiP cells, where the following analyses were carried out: D) immunoblotting for CAR and TRAP in post-production cell lysates and concentrated LV particles, E) vRNA splicing analysis of total cell RNA, and vector RNA by RT-PCR using primers #1 and #2, and F) efficiency of T-cell transduction by CAR/flow cytometry (% CAR-positive, n = 4) or vector copy-number analysis (VCN, n = 2) or percent target cell killing (n = 4) (ND – not detected). Data are mean [SD]; log10-transformed data.

within the pCMV-GFP reporter harbouring the Ψ-MSD-EF1αSA spliced leader. These data show that the presence of the Ψ-MSD-spliced 5′ leader within GOI-encoding mRNA was poorly repressible by TRAP irrespective of the proximity of the tbs to the Kozak, and that in the context of LV cassettes the improved TRiP 5′UTR-tbs sequences were only manifest when implemented within 2KO-LVs (Fig. 7C). This implies that the degree of aberrant MSD splicing within a 3rd generation LV can be a major determinant on TRiP repression, and therefore 2KO-LVs are greatly improved vehicles to implement the TRiP System™.

3.7. *Enhanced production and quality of 2KO-LVs encoding therapeutic genes using the TRiP system*

We applied the improved TRiP System™ for use in the production of therapeutic LVs by onboarding 2KO-LV genomes with two different chimeric antigen receptor (CAR) genes driven by the EF1α or EFS promoters. The first encoded a CAR targeted to the 5T4

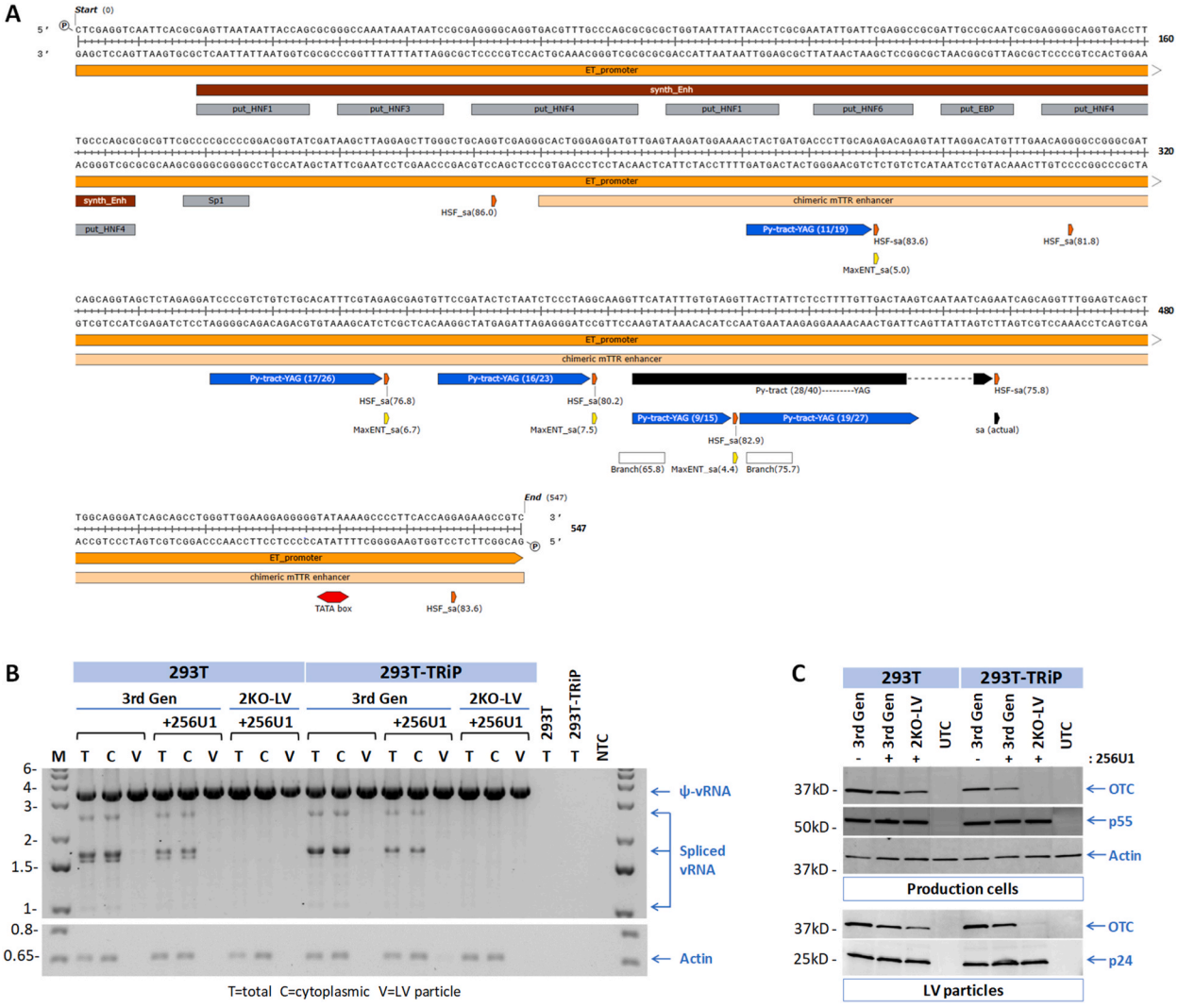


Fig. 9. Comparison of 3rd generation and 2KO-LV/TRiP vector genomes in production of LVs encoding the tissue-specific Enh1mTTR ('ET') promoter. Ornithine Transcarbamylase [OTC] encoded 3rd generation or 2KO-LV vectors were generated where the OTC gene was driven by the ET promoter; sequence shown in A. The synthetic enhancer region is composed of putative hepatocyte-specific transcription factor binding sites (grey boxes), which is fused to the chimeric murine transthyretin promoter (TATA box in red). The ET promoter was subject to in silico analysis, and the top hits for putative splice acceptors are annotated (HSF3.1 score x/100, MaxENT x/10). In addition, polypyrimidine tracts of the most obvious putative splice acceptors are highlighted in blue, with the actual polypyrimidine tract and splice acceptor site used by the MSD in black (potential branch sites in white boxes). LVs were produced in suspension HEK293T or HEK293T-TRAP cells, where the following analyses were carried out: B) vRNA splicing analysis of total or cytoplasmic cell RNA, and vector RNA by RT-PCR using primers #1 and #2, and C) immunoblotting for OTC and p55 (Gag) in post-production cell lysates and OTC and p24 (capsid) in concentrated LV particles. The major spliced vRNA product at ~1650bp detected during production of the 3rd generation LV was sequenced to identify the spliced acceptor site (sa actual) within the ET promoter to which the MSD spliced, despite this site having a low HSF3.1 score (75.8/100), and not appearing to be predicted by MaxENT.

protein, previously shown to be problematic during LV production, leading to low titres [23]. We generated 3rd generation and 2KO [m5]-LVs in suspension HEK293T-TRiP cells and analysed production cell surface expression of CAR (Fig. 8A), as well as performing RNAseq on total and unspliced vRNA in post-production cells and LV particles. These data demonstrate the absence of spliced vRNA within 2KO-LVs (Fig. 8B). Titration of clarified harvested LV showed that the 2KO-LVs were produced at >25-fold greater titres than the 3rd generation LV (Fig. 8C).

The second CAR was the CD19-FMC63 protein [38], which does not appear to impact on LV titres when produced in transient transfection. Nevertheless, we applied the TRiP System™ to these vectors to further demonstrate the capability of the TRiP System™ to greatly minimise the presence of CAR protein within both production cells and final drug product (Fig. 8D), as well as eliminate the presence of spliced vRNA (Fig. 8E). The 3rd generation and 2KO[m5] LV-CAR(CD19) vectors were able to transduce activated T-cells equivalently, leading to functional CAR-T cells (Fig. 8F).

Finally, we modelled the production of an LV-based product for liver-directed therapy (Ornithine Transcarbamylase [OTC] Deficiency) by generating LV genomes encoding codon-optimised OTC, driven by the semi-synthetic Enh1mTTR ('ET') promoter. This promoter contains hepatic transcription factor binding sites fused to the murine transthyretin promoter (Fig. 9A) and is expected to have low activity in HEK293(T) vector production cells [39]. Both 3rd generation and 2KO[m5] LVs were produced in suspension HEK293T or HEK293T-TRiP cells to generate LV stocks of similar output titres (within 2-fold, data not shown). Analysis of post-production cell RNA revealed spliced vRNA products were produced by the 3rd generation LV genome cassette, yielding a main spliced product of ~1.6 kb. Sequencing of this product revealed an MSD-splice acceptor junction within the ET core promoter element (Fig. 9A and B). The majority of OTC protein expressed during 3rd generation LV production - and detectable in resultant LV particles - appeared to be derived from this spliced vRNA product, since the 2KO-LV vector (no spliced vRNA; Fig. 9B) generated far less OTC protein in HEK293T cells (i.e. from the ET promoter only; Fig. 9C). Employment of the TRiP system™ was partially able to limit OTC production from the 3rd generation LV genome plasmid in HEK293T-TRiP cells but OTC production and incorporation into LV particles was only fully ablated when utilising the TRiP system™ within the 2KO-LV genome. Sequence analysis of the cryptic splice acceptor site within the ET promoter (Human Splicing Finder 3.1) did not provide a high predictive score of use of this site relative to other potential sites nearby. This implies that a general approach of removing/minimising aberrant splicing activity within 3rd generation LV genome cassettes by eliminating splice acceptor sites may be laborious and not necessarily successful, as well as being undesirable if changes to *cis*-acting sequences (such as promoters) are required.

4. Discussion

The development of lentiviral vectors as gene delivery vehicles has been extremely successful, with iterative improvements across three generations. The HIV-1 genome is complex, and its replication cycle depends on the plasticity of alternative splicing events to generate a balance of singly and multiply spliced, as well as unspliced, mRNAs at early and late times of infection. Our results reveal an imperfect configuration of typical contemporary lentiviral vector genomes, where the position of the RRE and use of Rev is expected to mediate the production of primarily full-length genomic RNA. This perhaps indicates an inherent structural preference for the MSD to 'find' strong consensus or cryptic splice acceptor sites that are central-distal to the RNA being produced, which is typical for the wild-type HIV-1 genome. In addition to recent publications reporting on LV vectorology [16,22], we are aware of this type of aberrant MSD splicing occurring in other available LV genome constructs, demonstrating that this effect is not peculiar to our LV system. This explains why tissue specific promoters with low/no activity within HEK293-based viral vector production cells appear (somewhat perplexingly) to become active when moving a transgene cassette from, for example, an rAAV genome plasmid to a lentiviral genome plasmid. Given that the major aberrantly spliced (and consequently rev/RRE-independent) mRNA products encode for the transgene ORF, and is therefore more likely to be 'seen' as a legitimate mRNA to be exported to the cytoplasm, the degree of total nuclear splicing from the MSD to generate products other than full length vRNA may be underestimated in our work. Our study is limited in that we have not attempted to uncover a quantitative/predictable mechanism by which transgene sequences are 'considered' by the splicing machinery as more or less favourable splice acceptor sites, when committed to splicing from the MSD. Such a model could aid payload sequence optimisation within MSD-active LVs to minimise aberrant splicing in LV production. However, the MSD could still be utilised within the integrated LV backbone by 5′ transcriptional read-through, and splicing out to adjacent cellular splice acceptors. The example of the EF1α promoter-intron sequence in our study appears to represent an extreme example of MSD mis-splicing, but the other (albeit non-exhaustive) examples generally argue for 3rd generation LV vector designs (i.e. containing an active MSD) that avoid transgene cassette introns entirely.

Here we have shown that MSD-spliced vRNA products can be translated efficiently during LV production, as well as detected within purified LV product, generally correlating with cellular abundance. Incorporation presumably occurs due to Ψ-independent packaging as virions bud from the cells, just as any other abundant transcript or protein, since these spliced vRNAs lack the core ψ-element [40,41]. Whilst spliced vRNAs also lack the central polypurine tract (cppt), and therefore are likely to be reverse-transcribed less efficiently, others have shown that HIV-1 mRNAs of similar architecture to these spliced vRNAs can be reverse-transcribed [42,43]. We are not aware of any literature that has reported the detection of shortened DNA cassettes integrated into target cells after LV transduction. However, this probably reflects the relative paucity of data from high dose LV applications to date, as well as the infrequent use of long-read sequencing analyses. Here we have shown that indeed these spliced vRNAs are templates for cDNA synthesis, but the resultant products appear to be primarily episomal and are ultimately lost from dividing cells. Our data imply that the cDNA episomes resulting from MSD-spliced vRNA may account for most of the episomal cDNA 'counted' as 1–2 LTR circles or linear forms in previous literature [44], and warrants further investigation. As the revolution in LV targeting continues, the use of LVs for high dose *in vivo* use will encourage further scrutiny of their safety in the same manner currently being carried out in the rAAV field [3,6,45]. Therefore, the

ability to produce LVs with simplified and consistent vRNA profiles in the manner demonstrated here with 2KO-LVs, such that essentially no episomal cDNA by-products are produced in target cells, is a sensible precautionary step to this end.

An alternative solution is to invert the transgene cassette relative to the vRNA cassette so that any such spliced vRNA derived from the MSD are unable to encode the transgene, and less likely to be 'seen' as a legitimate mRNA for nuclear export. However, this will result in the production of double-stranded (ds)RNA, triggering of dsRNA responses and reduced LV output titres, unless the transgene promoter has very low activity. Whilst there are ways of abrogating dsRNA sensing pathways to rescue output titres [46–48], we instead pursued a simpler line of research to achieve complete MSD inactivation, as well as ablate the use of nearby cryptic splice donor sites in the core ψ region. Past evaluations of MSD-inactivation within different LV generation genomes produce different outcomes, and this is primarily due to whether the system employs Tat. It's clear from those observations (and in this report) that Tat negates the otherwise dominant effect of mutating the MSD that leads to lower vRNA production and output titres. However, the use of Tat is not desirable in LV production systems, and would be seen as a retrograde step in LV development. Mutation of the MSD likely leads to a loss of endogenous U1 snRNA recruitment to the emerging vRNA transcript. Whilst we initially theorised that this might lead to premature polyadenylation at the poly-A site at the 5′ R-U5 boundary (as per HIV-1), the mutation of the poly-A signal did not substantially increase MSD-mutated LV titres. This indicates that U1-telescripting at the 5′ LTR in LV genome cassettes is not a key feature shared with HIV-1. Given U1 snRNA is also recruited to the transcription machinery, stimulates elongation, and couples transcription with intron-splicing, it's conceivable that Tat and U1 snRNA recruitment have redundant functions in moving-on prematurely stalled transcription complexes [49,50]. The '2KOm5' variant was designed to increase potential base-paring with endogenous U1 snRNA without being a useable splice donor. The idea that the MSD region (SL2) and U1 snRNA recruitment play an important role in vRNA biogenesis and/or stability, in addition to alternative splicing, is supported by the phenotype of the 2KOm5 variant [51]. It gave rise to the highest LV titres, with or without re-targeted U1 snRNA and did not induce crSD3/4 splicing compared to other MSD-mutants. Others have suggested that SL2 containing the MSD has a role in vRNA packaging [52], however our additional packaging-competition studies with 2KOm5 and wild type ψ-LV vRNA show that both species package with equal efficiency (data not shown).

The re-targeted U1 snRNAs appear to mediate their greatest effect when targeted to the same core packaging signal region in both 3rd generation and MSD-mutated LVs. Due to the limited number of LV genomes used for screening optimal re-targeted U1 snRNA binding sites, we cannot exclude the possibility that alternative target sites (compared to 256U1's) across the 5′ packaging region might be slightly more optimal for LVs containing different transgene cassettes. Nevertheless, we consistently observed a similar magnitude of 2KO-LV titre rescue for different LV payloads when using 256U1. Interestingly, 256U1 snRNA binds across the dimerization sequence in SL1; vRNA dimerization is a key early step in virion assembly. This indicates that the enhancer effect and subsequent release of 256U1 snRNA from its target sequence occurs *before* any intra/intermolecular vRNA folding events that may be required for dimerization of vRNAs. The dependency of re-directed U1 snRNAs on the Sm-region shows that typical snRNA processing is required, and that they are likely to be encompassed within mature U1 snRNP complexes, being primarily nuclear. The magnitude of clearance of 256U1 snRNA signal through downstream processing is consistent with this signal being present as a host cell residual i.e. coming from fragile cells, rather than the 256U1 snRNA being remained bound to vRNA and co-packaged into vector particles. Endogenous U1 snRNA has been detected in HIV-1 preparations, and the signal of endogenous U1 snRNA compared to 256U1 snRNA within final LV drug product in this study was ~10-to-1. This demonstrates that the presence of the class of U1 snRNA within LV drug product is neither unexpected nor dominated by the re-targeted snRNA species [40]. In further studies, we have neither been able to detect 256U1 snRNA DNA sequences in LV-transduced T-cells, nor have found evidence of off-target effects on host cell RNA sequences that share ≥12 nucleotides of homology to the 256U1 snRNA target sequence (data not shown).

The consequence of aberrant splicing from the MSD on LV output titres appears to be product-specific. The actual vector titre achieved for any given LV product generated by transient transfection, where plasmid DNA molecules may number up to 50,000 per cell [53], can still be relatively high even though spliced vRNA may be substantially in-excess of unspliced, packageable vRNA. A contributing factor may be the degree to which the transgene protein impacts on LV titres in an acute manner. Here we performed studies with two different CAR proteins [1]; the CD19-CAR, which did not require transgene repression to achieve high titres, and [2] the 5T4-CAR, which not only benefitted from TRiP repression but whose titres were also boosted in the context of 3rd generation LVs when 256U1 snRNA was co-expressed. For the LV-5T4-CAR vector, this was probably due to the modest but consistent reduction in transgene expression (data not shown), as similarly observed for other 3rd generation LV genomes assessed with 256U1. It remains to be tested whether this reflects minimal (and difficult to measure) reduction in production of GOI-encoding spliced vRNAs, or some inhibitory effect of translatability of full length vRNA, downstream of the 256U1 interaction. Thus, the use of 256U1 snRNA can boost some 3rd generation LVs by 2- to 3-fold but further increases (up to 10-fold) can be realised for transgenes that profoundly affect some aspect of output titre. The 256U1 expression cassette is only 690 nucleotides, and we have subsequently generated LV genome plasmids with this cassette included *in cis*, thus eliminating sourcing and costs for an additional GMP plasmid.

Also of note is the reduced impact of MSD-mutation on EF1α promoter-containing 2KO-LV genomes, compared to other promoter cassettes. On further analysis we found that titres were partially rescued for EF1α promoter-containing 2KO-LV genomes due to splicing-out of the EF1α intron (data not shown). This discovery has led us to develop an alternative class of 2KO-LV genomes, whose production is completely independent of both re-targeted U1 snRNA and Rev, allowing for deletion of RRE to increase space for transgene sequences (reference [54] and manuscript in preparation).

Of interest is the fact that for the generation of stable LV producer cell lines (PrCLs), where the vRNA component is typically limiting, the number of integrated LV genome expression cassettes is much lower than copy-numbers in transient transfection [55]. This indicates that aberrant splicing from the MSD to internal sequences, resulting in more spliced vRNA, may be in part contributing to the apparent block of the development of PrCLs. More specifically the generation of LV-CAR PrCLs has been made more difficult due

to an increased metabolic burden and/or toxicity on the PrCLs due to the constitutive expression of the CAR. This may place selective pressure at the stable pool stages, resulting in a preponderance of low or less active LV genome cassette-containing clones at the point of screening. We are now successfully combining these technologies in generating LV-CAR PrCLs, being able to more efficiently suppress transgene expression using the TRiP system™ embedded within 2KO-LV cassettes, and use of stably integrated 256U1 snRNA cassettes.

The TRiP system™, being vector agnostic, is widely applicable to other viral vector systems such as rAAV and AdVs. This further demonstrates that transgene suppression may be a useful approach to standardising viral vector production as a pipeline platform, where only the vector components are expressed during production, thus saving process development time/costs. There are other good reasons to ensure that the transgene protein is absent from the LV drug product. For example, the undesirable contamination of human Factor VIII within LV drug product intended for systemic delivery has necessitated others to apply an additional purification step in downstream processing, due to secretion from production cells into the upstream feed material [56]. Here, we have demonstrated that 2KO-LVs employing the TRiP system™ greatly minimises the presence of CAR protein in LVs. We did not explicitly look at the potential for LVs coated with CAR protein to contribute to off-target (leukemic) B-cell transduction due to the surface expression of the CAR antigen. But the realisation of such an event might lead to internalisation of the surface antigen due to binding of the expressed CAR, and result in 'pseudo' antigen-negative relapse [57]. This risk may be greater for *in vivo* administration where leukemic B-cells may be more abundant, and warrants further investigation.

In summary, the generation of a new class of so-called '2KO-LVs' and novel ψ-targeted U1 snRNA enhancers allows the production of LV particles with simplified vRNA profiles absent of spliced vRNAs, and when employing the TRiP system™, an LV product with greatly diminished transgene protein content. These new tools, which form part of the TetraVecta™ System, can be used to realise greater LV yields depending on the transgene payload and increase the likelihood of successful PrCL development. These improvements to LV design will further contribute to ensuring the safety of LVs for the next wave of gene therapy products, particularly for high dose *in vivo* applications where LV product quality is likely to play a more important role compared to current *ex vivo* applications.

CRediT authorship contribution statement

J. Wright: Writing – review & editing, Supervision, Methodology, Investigation, Data curation, Conceptualization. **B.M. Alberts:** Writing – review & editing, Methodology, Investigation. **A.J.M. Hood:** Writing – review & editing, Methodology, Investigation. **C. Nogueira:** Writing – review & editing, Methodology, Investigation. **Z. Miskolczi:** Writing – review & editing, Methodology, Investigation. **C.R. Vieira:** Writing – review & editing, Methodology, Investigation. **D. Chipchase:** Writing – review & editing, Software, Methodology, Investigation, Data curation. **C.M. Lamont:** Writing – review & editing, Methodology, Investigation. **O. Goodyear:** Writing – review & editing, Methodology, Investigation. **L.J. Moyce:** Writing – review & editing, Methodology, Investigation. **M. Soyombo:** Writing – review & editing, Methodology, Investigation. **D. Blount:** Writing – review & editing, Methodology, Investigation. **A.L. Keating:** Writing – review & editing, Methodology, Investigation. **T. Coradin:** Writing – review & editing, Methodology, Investigation. **H. Huang:** Writing – review & editing, Methodology, Investigation. **M. Martin-Urdiroz:** Writing – review & editing, Methodology, Investigation. **S. Ferluga:** Writing – review & editing, Methodology, Investigation. **K.A. Mitrophanous:** Writing – review & editing, Resources, Funding acquisition, Conceptualization. **N.G. Clarkson:** Writing – review & editing, Resources, Funding acquisition, Conceptualization. **D.C. Farley:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Data curation, Conceptualization.

Data availability statement

Data will be made available on request. For requesting data, please write to the corresponding author (DC Farley).

Ethics

The work conducted did not involve animal studies or clinical data, and therefore did not require an approving body. It contains no patient data. The work therefore did not need/have an approval number.

Trial registration numbers (where appropriate)

Not applicable.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:J Wright reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. BM Alberts reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. C Nogueira reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. Z Miskolczi reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. CR Vieira reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. D Chipchase reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. CM Lamont reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. O Goodyear reports a

relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. LJ Moyce reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. M Soyombo reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. D Blount reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. AL Keating reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. T Coradin reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. H Huang reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. M Martin-Urdiroz reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. S Ferluga reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. KA Mitrophanous reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. NG Clarkson reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. DC Farley reports a relationship with Oxford Biomedica Plc that includes: employment and equity or stocks. J Wright has patent ##WO2021160993A1 pending to Oxford Biomedica. J Wright has patent ##WO2021014157A1 pending to Oxford Biomedica. DC Farley has patent ##WO2021160993A1 pending to Oxford Biomedica. DC Farley has patent ##WO2021014157A1 pending to Oxford Biomedica. DC Farley has patent ##WO2021094752A1 pending to Oxford Biomedica. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2025.e43732.

References

[1] L. Naldini, U. Blömer, P. Gallay, D. Ory, R. Mulligan, F.H. Gage, I.M. Verma, D. Trono, In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector, Science 272 (1996) 263–267, https://doi.org/10.1126/science.272.5259.263.

[2] W.J. Holst, M.J. Giehm, Delivering genes with human immunodeficiency virus-derived vehicles: still state-of-the-art after 25 years, J. Biomed. Sci. 29 (2022), https://doi.org/10.1186/s12929-022-00865-4.

[3] D.E. Sabatino, F.D. Bushman, R.J. Chandler, R.G. Crystal, B.L. Davidson, R. Dolmetsch, K.C. Eggan, G. Gao, I. Gil-Farina, M.A. Kay, D.M. McCarty, E. Montini, A. Ndu, J. Yuan, Evaluating the state of the science for adeno-associated virus integration: an integrated perspective, Mol. Ther. 30 (2022) 2646–2663, https://doi.org/10.1016/j.ymthe.2022.06.004.

[4] A. Rezza, C. Jacquet, A.L. Pillouer, F. Lafarguette, C. Ruptier, M. Billandon, P.I. Petit, S. Trouttet, K. Thiam, A. Fraichard, Y. Chérifi, Unexpected genomic rearrangements at targeted loci associated with CRISPR/Cas9-mediated knock-in, Sci. Rep. 9 (2019), https://doi.org/10.1038/s41598-019-40181-w.

[5] G. Burgio, L. Teboul, Anticipating and identifying collateral damage in genome editing, Trends Genet. 36 (2020) 905–914, https://doi.org/10.1016/j.tig.2020.09.011.

[6] Z. Wang, P.J. Troilo, T.G. Griffiths, L.B. Harper, A.B. Barnum, S.J. Pacchione, C.J. Pauley, J.A. Lebron, J. Wolf, B.J. Ledwith, Characterization of integration frequency and insertion sites of adenovirus DNA into mouse liver genomic DNA following intravenous injection, Gene Ther. 29 (2022) 322–332, https://doi.org/10.1038/s41434-021-00278-2.

[7] F.D. Bushman, DNA transposon mechanisms and pathways of genotoxicity, Mol. Ther. 31 (2023) 613–615, https://doi.org/10.1016/j.ymthe.2023.01.023.

[8] C. Goyvaerts, T. Liechtenstein, C. Bricogne, D. Escors, K. Breckpot, Targeted Lentiviral Vectors: Current Applications and Future Potential, Gene Therapy - Tools and Potential Applications, 2013, https://doi.org/10.5772/52770.

[9] N. Cordes, N. Winter, C. Kolbe, B. Kotter, J. Mittelstaet, M. Assenmacher, T. Cathomen, A. Kaiser, T. Schaser, Adapter-mediated transduction with lentiviral vectors: a novel tool for cell-type-specific gene transfer, Viruses 14 (2022) 2157, https://doi.org/10.3390/v14102157.

[10] E. Martínez-Molina, C. Chocarro-Wrona, D. Martínez-Moreno, J.A. Marchal, H. Boulaiz, Large-scale production of lentiviral vectors: current perspectives and challenges, Pharmaceutics 12 (2020) 1051, https://doi.org/10.3390/pharmaceutics12111051.

[11] B. Eric, Regulating the gene-therapy revolution, Nature 564 (2018) S20–S22, https://doi.org/10.1038/d41586-018-07641-1.

[12] D.S. Anson, M. Fuller, Rational development of a HIV-1 gene therapy vector, J. Gene Med. 5 (2003) 829–838, https://doi.org/10.1002/jgm.415.

[13] E. Clark, B. Nava, M. Caputi, Tat is a multifunctional viral protein that modulates cellular gene expression and functions, Oncotarget 8 (2017) 27569–27581, https://doi.org/10.18632/oncotarget.15174.

[14] K. Cornetta, J. Yao, A. Jasti, S. Koop, M. Douglas, D. Hsu, L.A. Couture, T. Hawkins, L. Duffy, Replication-competent lentivirus analysis of clinical grade vector products, Mol. Ther. 19 (2011) 557–566, https://doi.org/10.1038/mt.2010.278.

[15] F.D. C, Development of a replication-competent lentivirus assay for dendritic cell-targeting lentiviral vectors, Mol Ther Methods Clin Dev 2 (2015) 15017, https://doi.org/10.1038/mtm.2015.17.

[16] N.P. Sweeney, C.A. Vink, The impact of lentiviral vector genome size and producer cell genomic to gag-pol mRNA ratios on packaging efficiency and titre, Mol Ther Methods Clin Dev 21 (2021) 574–584, https://doi.org/10.1016/j.omtm.2021.04.007.

[17] A.R. Cooper, G.R. Lill, E.H. Gschweng, D.B. Kohn, Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter, Nucleic Acids Res. 43 (2015) 682–690, https://doi.org/10.1093/nar/gku1312.

[18] C. May, S. Rivella, J. Callegari, G. Heller, K.M.L. Gaensler, L. Luzzatto, M. Sadelain, Therapeutic haemoglobin synthesis in β-thalassaemic mice expressing lentivirus-encoded human β-globin, Nature 406 (2000) 82–86, https://doi.org/10.1038/35017565.

[19] Y. Cui, T. Iwakuma, L.-J. Chang, Contributions of viral splice sites and cis-Regulatory elements to lentivirus vector function, J. Virol. 73 (1999) 6171–6176, https://doi.org/10.1128/JVI.73.7.6171-6176.1999.

[20] M.R. Mautino, W.J. Ramsey, J. Reiser, R.A. Morgan, Modified human immunodeficiency virus-based lentiviral vectors display decreased sensitivity to trans-dominant, rev. Hum. Gene Ther. 11 (2000) 895–908, https://doi.org/10.1089/10430340050015509.

[21] A. Schambach, D. Mueller, M. Galla, M.M.A. Verstegen, G. Wagemaker, R. Loew, C. Baum, J. Bohne, Overcoming promoter competition in packaging cells improves production of self-inactivating retroviral vectors, Gene Ther. 13 (2006) 1524–1533, https://doi.org/10.1038/sj.gt.3302807.

[22] H. Sertkaya, M. Ficarelli, N.P. Sweeney, H. Parker, C.A. Vink, C.M. Swanson, HIV-1 sequences in lentiviral vector genomes can be substantially reduced without compromising transduction efficiency, Sci. Rep. 11 (2021), https://doi.org/10.1038/s41598-021-91309-w.

[23] H.E. Maunder, J. Wright, B.R. Kolli, C.R. Vieira, T.T. Mkandawire, S. Tatoris, V. Kennedy, S. Iqball, G. Devarajan, S. Ellis, Y. Lad, N.G. Clarkson, K. A. Mitrophanous, D.C. Farley, Enhancing titres of therapeutic viral vectors using the transgene repression in vector production (TRiP) system, Nat. Commun. 8 (2017), https://doi.org/10.1038/ncomms14834.

[24] D. Farley, Standardizing viral vector manufacture: maximizing production with the TRiP SystemTM, Cell Gene Ther. Insights 4 (2018) 983–996, https://doi.org/10.18609/cgti.2018.099.

[25] Y. Ran, Y. Deng, C. Yao, U1 snRNP telescripting: molecular mechanisms and beyond, RNA Biol. 18 (2021) 1512–1523, https://doi.org/10.1080/15476286.2021.1872963.
[26] L. Blázquez, P. Fortes, U1 interference (U1i) for Antiviral approaches, Adv. Exp. Med. Biol. (2015), https://doi.org/10.1007/978-1-4939-2432-5_3.
[27] K.R. M, Refinement of lentiviral vector for improved RNA processing and reduced rates of self inactivation repair, BMC Biotechnol. 9 (2009) 86, https://doi.org/10.1186/1472-6750-9-86.
[28] Mark P. Ashe, Louise H. Pearson, Nick J. Proudfoot, The HIV-1 5′ LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site, EMBO J. 16 (1997) 5752–5763, https://doi.org/10.1093/emboj/16.18.5752.
[29] B. Jens, Mutation of the major 5′ splice site renders a CMV-driven HIV-1 proviral clone tat-dependent: connections between transcription and splicing, FEBS Lett. 563 (2004) 113–118, https://doi.org/10.1016/S0014-5793(04)00277-7.
[30] A.M. P., Stem-loop 1 of the U1 snRNP plays a critical role in the suppression of HIV-1 polyadenylation, RNA 6 (2000) 170–177, https://doi.org/10.1017/S1355838200991957.
[31] D.C. Farley, J. Wright, Improved SIN-LTRs for Lentiviral Vectors (Patent WO2023062365A2), n.d.
[32] K. Siarhei, Transcriptional start site heterogeneity modulates the structure and function of the HIV-1 genome, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 13378–13383, https://doi.org/10.1073/pnas.1616627113.
[33] X. Heng, S. Kharytonchyk, E.L. Garcia, K. Lu, S.S. Divakaruni, C. LaCotti, K. Edme, A. Telesnitsky, M.F. Summers, Identification of a minimal region of the HIV-1 5′-Leader required for RNA dimerization, NC binding, and packaging, J. Mol. Biol. 417 (2012) 224–239, https://doi.org/10.1016/j.jmb.2012.01.033.
[34] T. Jacks, M.D. Power, F.R. Masiarz, P.A. Luciw, P.J. Barr, H.E. Varmus, Characterization of ribosomal frameshifting in HIV-1 gag-pol expression, Nature 331 (1988) 280–283, https://doi.org/10.1038/331280a0.
[35] P. Rattananon, U. Anurathapan, K. Bhukhai, S. Hongeng, The future of gene therapy for transfusion-dependent beta-Thalassemia: the power of the lentiviral vector for genetically modified hematopoietic stem cells, Front. Pharmacol. 12 (2021), https://doi.org/10.3389/fphar.2021.730873.
[36] H. Li, Y. Lu, R.P. Witek, L.-J. Chang, M. Campbell-Thompson, M. Jorgensen, B. Petersen, S. Song, Ex vivo transduction and transplantation of bone marrow cells for liver gene delivery of α1-Antitrypsin, Mol. Ther. 18 (2010) 1553–1558, https://doi.org/10.1038/mt.2010.116.
[37] D.C. Farley, Transgene Repression (Vector) Production System (Patent WO2021094752A1), n.d.
[38] J. Seigner, C.U. Zajc, S. Dötsch, C. Eigner, E. Laurent, D.H. Busch, M. Lehner, M.W. Traxlmayr, Solving the mystery of the FMC63-CD19 affinity, Sci. Rep. 13 (2023), https://doi.org/10.1038/s41598-023-48528-0.
[39] E. Vigna, M. Amendola, F. Benedicenti, A.D. Simmons, A. Follenzi, L. Naldini, Efficient tet-dependent expression of human factor IX in vivo by a new self-regulating lentiviral vector, Mol. Ther. 11 (2005) 763–775, https://doi.org/10.1016/j.ymthe.2004.11.017.
[40] M.J. Eckwahl, H. Arnion, S. Kharytonchyk, T. Zang, P.D. Bieniasz, A. Telesnitsky, S.L. Wolin, Analysis of the human immunodeficiency virus-1 RNA packageome, RNA 22 (2016) 1228–1238, https://doi.org/10.1261/rna.057299.116.
[41] S. Johnson, J.X. Wheeler, R. Thorpe, M. Collins, Y. Takeuchi, Y. Zhao, Mass spectrometry analysis reveals differences in the host cell protein species found in pseudotyped lentiviral vectors, Biologicals 52 (2018) 59–66, https://doi.org/10.1016/j.biologicals.2017.12.005.
[42] L. Houzet, J.C. Paillart, F. Smagulova, S. Maurel, Z. Morichaud, R. Marquet, M. Mougel, HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms, Nucleic Acids Res. Suppl. 35 (2007) 2695–2704, https://doi.org/10.1093/nar/gkm153.
[43] L. Houzet, Z. Morichaud, M. Mougel, Fully-spliced HIV-1 RNAs are reverse transcribed with similar efficiencies as the genomic RNA in virions and cells, but more efficiently in AZT-treated cells, Retrovirology 4 (2007), https://doi.org/10.1186/1742-4690-4-30.
[44] C. Iglesias, M. Ringeard, F.D. Nunzio, J. Fernandez, R. Gaudin, P. Souque, P. Charneau, N. Arhel, Residual HIV-1 DNA Flap-independent nuclear import of cPPT/CTS double mutant viruses does not support spreading infection, Retrovirology 8 (2011), https://doi.org/10.1186/1742-4690-8-92.
[45] D.R. Deyle, D.W. Russell, Adeno-associated virus vector integration, Curr. Opin. Mol. Therapeut. 11 (2009) 442–447.
[46] B.C. Poling, K. Tsai, D. Kang, L. Ren, E.M. Kennedy, B.R. Cullen, A lentiviral vector bearing a reverse intron demonstrates superior expression of both proteins and microRNAs, RNA Biol. 14 (2017) 1570–1579, https://doi.org/10.1080/15476286.2017.1334755.
[47] H. Peirong, B. Yanmin, M. Hong, S. Thipparat, Z. Brian, F.N. J, K.D. B, K. Tal, Superior lentiviral vectors designed for BSL-0 environment abolish vector mobilization, Gene Ther. 25 (2018) 454–472, https://doi.org/10.1038/s41434-018-0039-2.
[48] T. Maetzig, M. Galla, M.H. Brugman, R. Loew, C. Baum, A. Schambach, Mechanisms controlling titer and expression of bidirectional lentiviral and gammaretroviral vectors, Gene Ther. 17 (2010) 400–411, https://doi.org/10.1038/gt.2009.129.
[49] A. Furger, J.M.O.A. Binnie, B.A. Lee, N.J. Proudfoot, Promoter proximal splice sites enhance transcription, Genes Dev. 16 (2002) 2792–2799, https://doi.org/10.1101/gad.983602.
[50] Y. Leader, G.L. Maor, M. Sorek, R. Shayevitch, M. Hussein, O. Hameiri, L. Tammer, J. Zonszain, I. Keydar, D. Hollander, E. Meshorer, G. Ast, The upstream 5′ splice site remains associated to the transcription machinery during intron synthesis, Nat. Commun. 12 (2021), https://doi.org/10.1038/s41467-021-24774-6.
[51] S. Kammler, C. Leurs, M. Freund, J. Krummheuer, K. Seidel, T.Ø. Tange, M.K. Lund, J. Kjems, A. Scheid, H. Schaal, The sequence complementarity between HIV-1 5′ splice site SD4 and U1 snRNA determines the steady-state level of an unstable env pre-mRNA, RNA 7 (2001) 421–434, https://doi.org/10.1017/S1355838201001212.
[52] J. Deforges, N. Chamond, B. Sargueil, Structural investigation of HIV-1 genomic RNA dimerization process reveals a role for the major Splice-site donor stem loop, Biochimie 94 (2012) 1481–1489, https://doi.org/10.1016/j.biochi.2012.02.009.
[53] R.N. Cohen, M.A.E.M. Aa, N. Macaraeg, A.P. Lee, F.C. Szoka, Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection, J. Contr. Release 135 (2009) 166–174, https://doi.org/10.1016/j.jconrel.2008.12.016.
[54] D.C. Farley, J. Wright, Rev/RRE-independent Lentiviral Vectors (Patent WO2023062363A1), n.d.
[55] R.E. Throm, A.A. Ouma, S. Zhou, A. Chandrasekaran, T. Lockey, M. Greene, S.S.D. Ravin, M. Moayeri, H.L. Malech, B.P. Sorrentino, J.T. Gray, Efficient construction of producer cell lines for a SIN lentiviral vector for SCID-X1 gene therapy by concatemeric array transfection, Blood 113 (2009) 5104–5110, https://doi.org/10.1182/blood-2008-11-191049.
[56] Ismail, A., Liu, T., Mayani, M, Methods for the Removal of Free Factor VIII from Preparations of Lentiviral Vectors Modified to Express Said Protein (Patent WO2021262963A1), n.d.
[57] M. Ruella, J. Xu, D.M. Barrett, J.A. Fraietta, T.J. Reich, D.E. Ambrose, M. Klichinsky, O. Shestova, P.R. Patel, I. Kulikovskaya, F. Nazimuddin, V.G. Bhoj, E. J. Orlando, T.J. Fry, H. Bitter, S.L. Maude, B.L. Levine, C.L. Nobles, F.D. Bushman, R.M. Young, J. Scholler, S.I. Gill, C.H. June, S.A. Grupp, S.F. Lacey, J. J. Melenhorst, Induction of resistance to chimeric antigen receptor T cell therapy by transduction of a single leukemic B cell, Nat. Med. 24 (2018) 1499–1503, https://doi.org/10.1038/s41591-018-0201-9.
[58] V.N. Kim, K. Mitrophanous, S.M. Kingsman, A.J. Kingsman, Minimal requirement for a lentivirus vector based on human immunodeficiency virus type 1, J. Virol. 72 (1998) 811–816, https://doi.org/10.1128/JVI.72.1.811-816.1998.
[59] E. Kotsopoulou, V.N. Kim, A.J. Kingsman, S.M. Kingsman, K.A. Mitrophanous, A rev-independent human immunodeficiency virus type 1 (HIV-1)-Based vector that exploits a codon-optimized HIV-1 GagPol gene, J. Virol. 74 (2000) 4839–4852, https://doi.org/10.1128/jvi.74.10.4839-4852.2000.
[60] D.C. Farley, S. Iqball, J.C. Smith, J.E. Miskin, S.M. Kingsman, K.A. Mitrophanous, Factors that influence VSV-G pseudotyping and transduction efficiency of lentiviral vectors - in vitro and in vivo implications, J. Gene Med. 9 (2007) 345–356, https://doi.org/10.1002/jgm.1022.
[61] G. Pamenter, L. Davies, C. Lamont, D. Rahim, C. Knevelman, J. Miskin, K. Mitrophanous, D. Dikicioglu, D.G. Bracewell, Lentiviral vector determinants of anion-exchange chromatography elution heterogeneity, Biotechnol. Bioeng. Symp. 121 (2024) 2936–2951, https://doi.org/10.1002/bit.28766.