Molecular Therapy
Methods & Clinical Development
Original Article

LPA disruption with AAV-CRISPR potently lowers plasma apo(a) in transgenic mouse model: A proof-of-concept study

Alexandria M. Doerfler,[1,7] So Hyun Park,[2,7] Julia M. Assini,[3] Amer Youssef,[4] Lavanya Saxena,[2] Adam B. Yaseen,[2] Marco De Giorgi,[1] Marcel Chuecos,[1] Ayrea E. Hurley,[1] Ang Li,[1,2] Santica M. Marcovina,[5] Gang Bao,[2] Michael B. Boffa,[3,4] Marlys L. Koschinsky,[4,6] and William R. Lagor[1,2]

[1]Department of Integrative Physiology, Baylor College of Medicine, Houston, TX 77030, USA; [2]Department of Bioengineering, Rice University, Houston, TX 77030, USA; [3]Department of Biochemistry, Schulich School of Medicine and Dentistry, the University of Western Ontario, London, ON N6A 5B7, Canada; [4]Robarts Research Institute, Schulich School of Medicine and Dentistry, London, ON N6G 2V4, Canada; [5]Medpace Reference Laboratories, Cincinnati, OH 45227, USA; [6]Department of Physiology and Pharmacology, Schulich School of Medicine and Dentistry, London, ON N6A 5B7, Canada

Lipoprotein(a) (Lp(a)) represents a unique subclass of circulating lipoprotein particles and consists of an apolipoprotein(a) (apo(a)) molecule covalently bound to apolipoprotein B-100. The metabolism of Lp(a) particles is distinct from that of low-density lipoprotein (LDL) cholesterol, and currently approved lipid-lowering drugs do not provide substantial reductions in Lp(a), a causal risk factor for cardiovascular disease. Somatic genome editing has the potential to be a one-time therapy for individuals with extremely high Lp(a). We generated an *LPA* transgenic mouse model expressing apo(a) of physiologically relevant size. Adeno-associated virus (AAV) vector delivery of CRISPR-Cas9 was used to disrupt the *LPA* transgene in the liver. AAV-CRISPR nearly completely eliminated apo(a) from the circulation within a week. We performed genome-wide off-target assays to determine the specificity of CRISPR-Cas9 editing within the context of the human genome. Interestingly, we identified intrachromosomal rearrangements within the *LPA* cDNA in the transgenic mice as well as in the *LPA* gene in HEK293T cells, due to the repetitive sequences within *LPA* itself and neighboring pseudogenes. This proof-of-concept study establishes the feasibility of using CRISPR-Cas9 to disrupt *LPA in vivo*, and highlights the importance of examining the diverse consequences of CRISPR cutting within repetitive loci and in the genome globally.

INTRODUCTION

Lipoprotein(a) (Lp(a)) is a lipoprotein particle that is well established as an independent and causal risk factor for atherosclerotic cardiovascular disease (ASCVD) and calcific aortic valve disease (CAVD).[1–9] Lp(a) is composed of a low-density lipoprotein (LDL)-like particle with apolipoprotein(a) (apo(a)) covalently linked to apolipoprotein B-100 (apoB-100).[10] The apo(a) protein itself is encoded by the *LPA* gene and is highly similar to plasminogen (*PLG*), a zymogen that plays a key role in fibrinolysis.[11] Both apo(a) and plasminogen contain a series of kringle (K) domains, which are triple loop structures stabilized by three internal disulfide bonds. The apo(a) protein consists of 10 kringle type IV (KIV) domains, denoted as KIV_1–KIV_{10}, followed by KV and protease-like domains. Lp(a) concentrations are strongly heritable and vary widely in humans based on the number of KIV_2 domains (ranging from two to greater than 40) that correspond to allele size of the *LPA* gene. Smaller isoforms of *LPA* encode apo(a) proteins that are more efficiently secreted from the liver, resulting in higher Lp(a) levels in the circulation.[12] Individuals with plasma Lp(a) levels over 30–50 mg/dL, or about one in five individuals in the United States,[13] are at increased risk for CAVD,[4,14] making it an important target for drug development.

Current lipid-lowering therapies, including PCSK9 inhibitors, do not target Lp(a) specifically or adequately lower plasma Lp(a) concentrations. RNA interference and antisense technologies targeting the *LPA* transcript show tremendous promise for reducing Lp(a) levels in high-risk patients.[15] Currently, there is an ongoing phase III clinical trial to assess whether antisense-mediated knockdown of apo(a) production by the liver can reduce the risk of major adverse cardiovascular events in patients with established cardiovascular disease (CVD)) and elevated Lp(a) levels (Lp(a) HORIZON, NCT04043552).[16] There are also small interfering RNA (siRNA) molecules in development that can effectively lower Lp(a), with a recent phase I study (NCT03626662) showing 71%–97% reductions after a single dose.[17] However, there is a strong case to be made for genome editing to achieve Lp(a) reduction. First, the *LPA* gene is exclusively expressed in the liver, an organ that is amenable to targeting with both viral

Received 25 May 2022; accepted 12 October 2022;
https://doi.org/10.1016/j.omtm.2022.10.009.

[7]These authors contributed equally

Correspondence: William R. Lagor, Department of Integrative Physiology, Baylor College of Medicine, Mail stop BCM 335, Houston, TX 77030, USA.
E-mail: lagor@bcm.edu

vectors and non-viral delivery systems. Second, although the biological function of apo(a) remains unknown, there appear to be no adverse phenotypes associated with homozygous loss of this gene in humans.[16,18,19] Third, oligonucleotides and siRNAs require repeated injections, whereas a single treatment with genome-editing enzymes could permanently eliminate apo(a) production, providing the greatest degree of durable CVD risk reduction without a need for additional dosing.

Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR-Cas9) is the most widely used system for *in vivo* gene editing, with many potential applications for human therapeutics.[20,21] Off-target effects remain a primary concern in the use of CRISPR-Cas9 for gene therapy as they can result in impaired cellular viability or tumorigenesis. Therefore, rigorous selection of single guide RNA (sgRNA) is required to reduce off-target effects and maximize editing efficiency. In addition, highly efficient delivery is a critical factor to maximize editing efficiency. Adeno-associated virus (AAV) has a high tropism for the liver and is a promising delivery vehicle for cardiovascular targets in the liver. AAV delivery of CRISPR-Cas9 has already been used successfully to edit *Pcsk9*, *Apob*, and *Ldlr*.[22–25] At the present time, preclinical gene editing to disrupt *LPA* has not been reported.

In this study, we utilized a gRNA specific for the first kringle in apo(a), and used AAV for delivery of the gRNA and Cas9. We measured the effect of editing on plasma apo(a) concentrations, and we documented on- and off-target effects on a genome-wide basis. This work demonstrates that liver-directed genome editing with CRISPR-Cas9 can effectively lower apo(a) levels, and highlights the importance of surveillance for unintended editing events within highly repetitive genes and globally within the genome.

RESULTS

Generation of transgenic mice expressing apo(a)

The *LPA* gene is found only in humans and a subset of primates, although a similar protein exists in hedgehogs that arose independently through a gene duplication event.[26] Therefore, we developed transgenic mice expressing a small version of human apo(a) by pronuclear injection of a linearized *LPA* cDNA driven by the liver-specific promoter from *APOE* (Figure 1). The *LPA* transgene was bred onto an Ldlr-null background to more closely reflect a human lipoprotein distribution with high levels of LDL. We found using droplet digital PCR (ddPCR) with transgene-specific primers that $LPA^{+/0}Ldlr^{-/-}$ mice have three *LPA* transgene copies per diploid genome (c/dg) (Figure S1). Sanger sequencing identified the junction containing contiguous sequence between an intergenic site on chr3:120164677 in the mouse genome (mm10) (Figure S2A) and the 5′ end of the *LPA* construct, with two types of the tandem junctions of the 3′ and 5′ ends of the *LPA* construct (Figures S2B and S2C). Therefore, the transgenic mice have three tandem repeats of the *LPA* construct in direct orientation integrated into the chr3:120164677 (Figure S2D). The original transgene encoded a 14-kringle form of apo(a), containing all the invariant KIV types and five copies of KIV_2. However, Sanger sequencing of the integrated transgene revealed the presence of an in-frame 342-bp deletion of most of KIV_8 and a small amount of KIV_7; the junction point was at the homologous site in the respective kringles and thus an intact hybrid of KIV_7 and KIV_8 was formed. Since no other alternative sequences were observed, the deletion occurred prior to tandem multiplication of the transgene insertion.

Characterization of Lp(a) transgenic mice

The *LPA* transgene was maintained in the hemizygous or null state for all experimental animals ($LPA^{+/0}$, $LPA^{+/0}Ldlr^{-/-}$, or $LPA^{0/0}Ldlr^{-/-}$) (Figure 1A). *LPA* transgene mRNA abundance in the liver, kidney, spleen, heart, small intestine, brain, and quadriceps was assessed by real-time RT-qPCR relative to *Gapdh* (Figure 1B). The highest levels of *LPA* expression were detected in liver and brain tissue, consistent with the *APOE* promoter elements driving the transgene. We performed a western blot to determine if apo(a) is present in the plasma and covalently bound with murine apoB-100 (Figure 1C). Human plasma expressing a form of Lp(a) that has 16 KIV domains (16K) was used as a control (C, lanes 1 and 10). Plasma from male and female mice lacking (0/0) or expressing (+/0) the human *LPA* transgene were resolved under reduced and non-reduced conditions. We observed that the apo(a) protein is detectable in plasma from both male and female mice but does not form a covalent disulfide bond with murine apoB-100 (Figure 1C) because the expected band of slightly higher mobility than the human Lp(a) control is absent. Apo(a) distribution on lipoproteins was assessed in lipoprotein fractions separated by size-exclusion chromatography from male mice that were fasted for 5 h. Cholesterol was distributed as expected within the profile for $Ldlr^{-/-}$ mice on a chow diet, and did not differ with respect to *LPA* genotype. Interestingly, apo(a) was found only in the lipoprotein fractions corresponding to very-low-density lipoprotein (VLDL)- and intermediate-density lipoprotein (IDL)/LDL-sized particles, with none detectable in high-density lipoprotein (HDL) or the lipoprotein free fractions (Figure 1D). Despite the inability to form a covalent disulfide bond with murine apoB-100, apo(a) still resides on apoB-100-containing particles in this mouse model, presumably through non-covalent interactions.

LPA gene editing rapidly and efficiently lowers circulating apo(a) protein levels

We generated an AAV vector co-expressing a sgRNA targeting *LPA* exon 2 in the KIV_1-encoding region of the *LPA* cDNA transgene. Given the highly repetitive nature of KIV domains in the *LPA* gene, and the need to achieve complete removal of apo(a) protein, we aimed to target as close as possible to the translation start site in the coding sequence. Exon 1 encodes only the first 49 bp of the signal peptide, and was not targetable by SaCas9. Multiple gRNAs were designed for exon 2 based on the presence of a canonical NNGRRT protospacer adjacent motif (PAM). The optimal gRNA for this region was found to reside in antisense orientation from +147 to +124 relative to the ATG start codon, within the first KIV domain. The designed 23-nucleotide sgRNA (GCAGGTCCTTCCTGTGACAGTGG, *PAM*-TGGAGT)

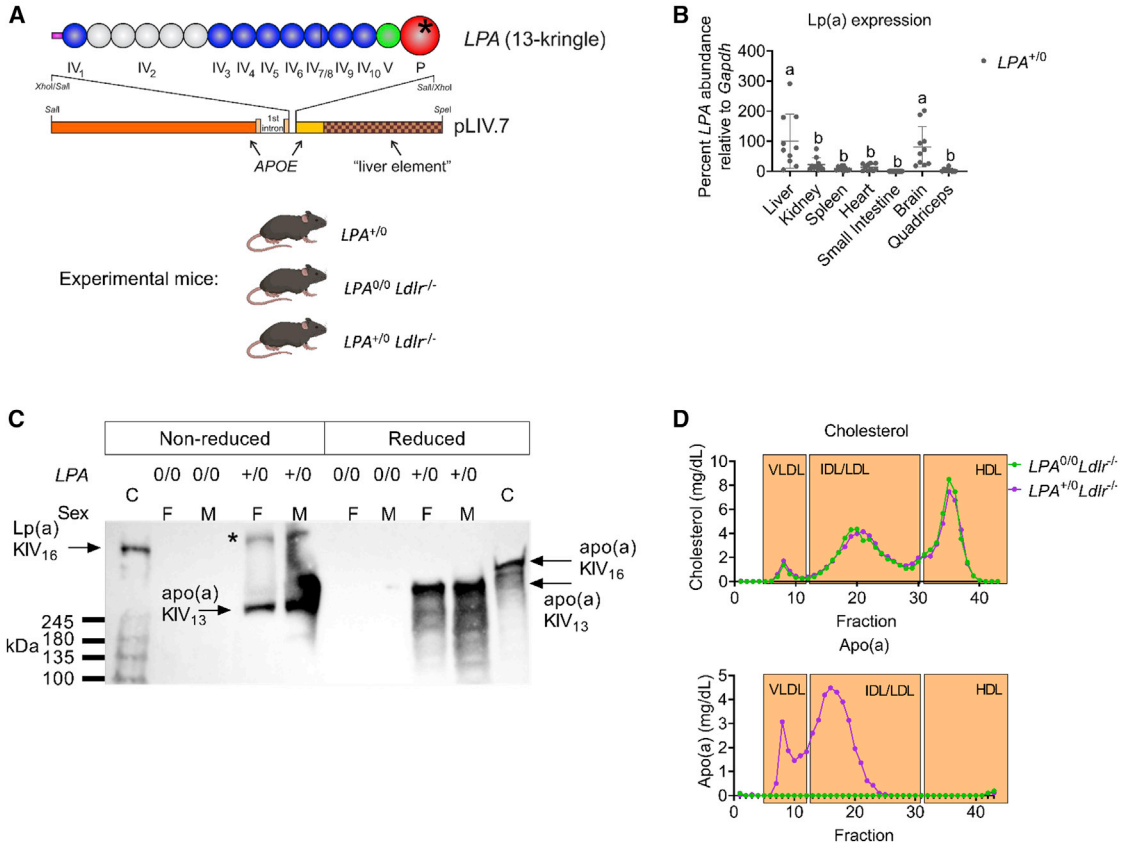


Figure 1. Generation of transgenic mice expressing apo(a)

(A) Schematic representation of the *LPA* transgene used to generate transgenic mice expressing human *LPA*. $LPA^{+/0}$ mice were then crossed with mice lacking the LDL receptor ($Ldlr^{-/-}$). (B) Relative *LPA* mRNA abundance in different $LPA^{+/0}$ mouse tissues as measured by qPCR (n = 10). Different letters denote data that are significantly different by ANOVA with *post hoc* Tukey test, $p < 0.05$. Values are mean ± SD. (C) Western blot for apo(a) in mice either lacking (0/0) or expressing (+/0) the *LPA* transgene, run under non-reduced or reduced conditions. M, male; F, female; C, control human plasma. *Aggregated material that did not enter the resolving gel. (D) Cholesterol and apo(a) distribution among lipoprotein fractions in pooled plasma (n = 4 per group).

had no exact-match off-targets identified in the *LPA* transgene. Expression of the sgRNA was driven by a U6 promoter, along with SaCas9 expressed under the control of the liver-specific hybrid liver promoter (HLP). Male and female $LPA^{+/0}Ldlr^{-/-}$ mice were injected intraperitoneally at 8 weeks of age with 1×10^{12} genome copies (GC) of AAV expressing either GFP as a control (AAV-GFP) or sgRNA and Cas9 (AAV-CRISPR). Plasma and body weights were collected at weekly intervals until the terminal time point at 12 weeks of age when livers were harvested (Figure 2A). There were no significant differences in body weights following treatment in either male or female mice throughout the study (Figure 2B). Liver to body weight ratios were modestly but significantly higher in male $LPA^{+/0}Ldlr^{-/-}$ mice that received AAV-CRISPR, while female ratios were unchanged between groups (Figure 2C). Cholesterol levels also remained unchanged between groups in male and female mice (Figure 2D). Liver histology for proliferation (Ki-67) and cell death (TUNEL) was quantified, and no significant changes were observed in mice that received AAV-CRISPR (Figure S3). Alanine transaminase (ALT) and aspartate transaminase (AST) levels were not significantly altered by AAV treatment, and remained in a normal range throughout the study (Figure S4). Finally, vector genome analysis by qPCR revealed approximately one to nine genome copies per diploid genome equivalent (dge) in mice that received AAV-GFP, while mice that received AAV-CRISPR ranged from one to 39 AAV genome copies per diploid genome (Figure S5). Although not significant, there is a trend toward lower AAV GC/dge in female mice, in accordance with previous studies that have shown reduced transduction in livers of female mice receiving AAV.[27] Robust expression of GFP and SaCas9 protein was confirmed by western blotting in mice receiving the respective AAV vectors (Figures 2E–2H). Plasma apo(a) protein expression was assessed by western blot, revealing undetectable levels in male and female $LPA^{+/0}Ldlr^{-/-}$ mice that received AAV-CRISPR at 4 weeks post injection (Figure 2I). An ELISA specific for apo(a) revealed essentially complete elimination of the protein from plasma as early as 1 week post injection, which was sustained for the duration of the study (Figure 2J). We also observed that apo(a) levels are approximately 2-fold higher in male transgenic mice relative to females at baseline.

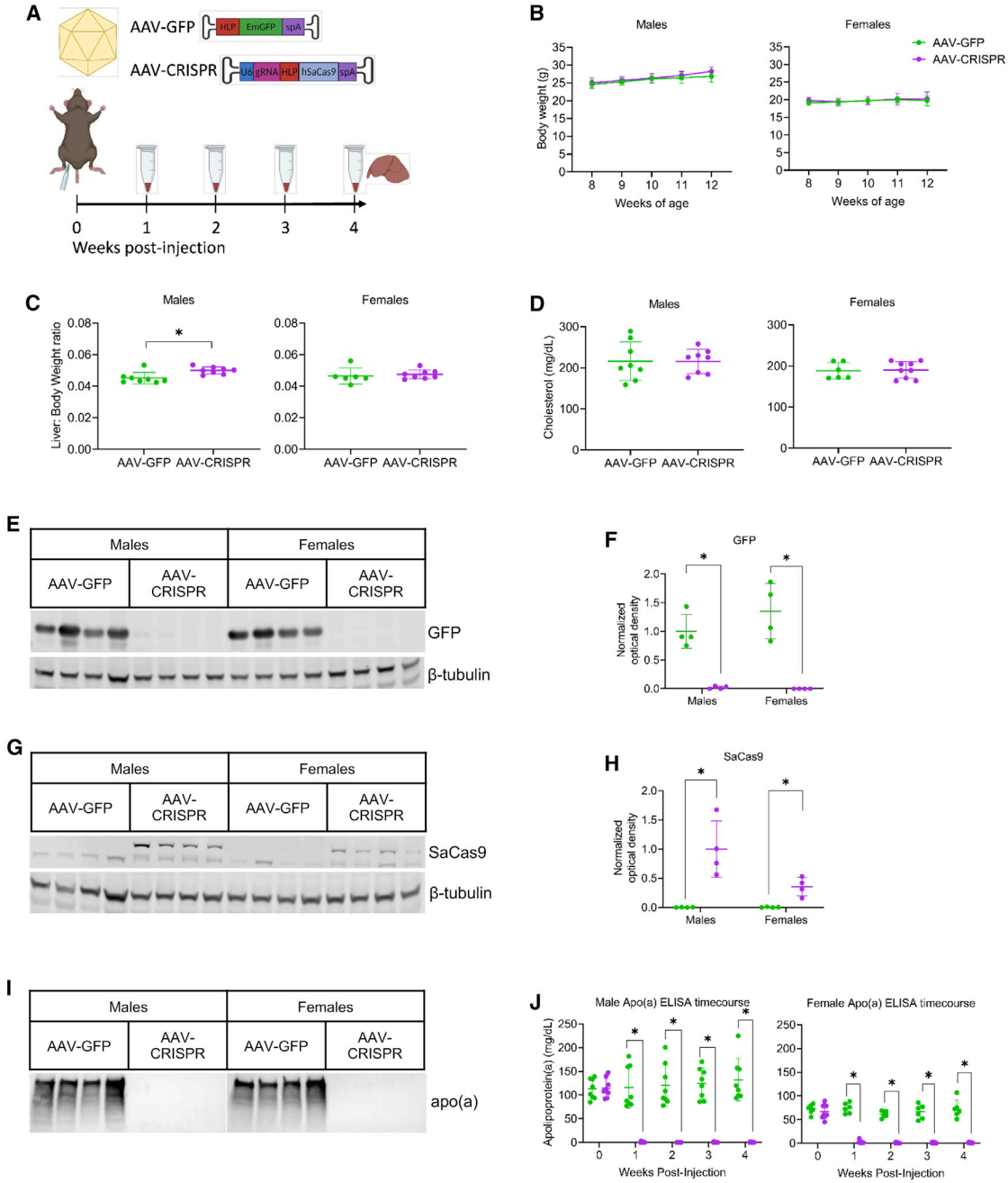


Figure 2. Knockdown of apo(a) in *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ mice by AAV-CRISPR

(A) Male and female *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ mice were injected at 8 weeks of age with 1×10^{12} genome copies of AAV expressing either GFP as a control (AAV-GFP) or Cas9 and an sgRNA targeting *LPA* exon 2 in the KIV_1 domain of the *LPA* transgene (AAV-CRISPR). Plasma and body weights were collected weekly until 12 weeks of age, at which point plasma and liver tissue were harvested. (B) Weekly body weights in male and female mice. (C) Liver to body weight ratios and (D) fasted plasma cholesterol levels in male and female mice given either AAV-GFP or AAV-CRISPR. (E) Western blot for GFP in male and female *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ mice given AAV-GFP or AAV-CRISPR. (F) Quantification for GFP western blot. All groups normalized to male mice given AAV-GFP vector. (G) Western blot for SaCas9 in male and female *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ mice given AAV-GFP or AAV-CRISPR. (H) Quantification for SaCas9 western blot. All groups normalized to male mice given AAV-GFP vector. (I) Western blot for apo(a) in male and female *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ mice given AAV-GFP or AAV-CRISPR. (J) Time course for apo(a) expression in male and female *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ mice given AAV-GFP or AAV-CRISPR measured by ELISA ($n \geq 6$). Statistical analysis was done using Mann-Whitney test on non-normally distributed data with *$p < 0.05$ (GFP western blot males, apo(a) ELISA males and females). Statistical analysis was done using a Welch's t test on normally distributed data with *$p < 0.05$ (body weight male and females at each time point, liver:body weight ratio males and females, cholesterol males and females, GFP western blot females, SaCas9 western blot males and females). Values are mean ± SD.

Assessment of gene editing in mice receiving AAV-CRISPR

We first measured editing at the on-target site in KIV_1, and assessed insertions and deletions (indels) by T7 endonuclease I (T7E1) assay and inference of CRISPR edits (ICE; data not shown) (Figures S6A and S6B). Both assays showed no measurable indels in all treated animals compared with control, which was surprising given the near-complete degree of plasma apo(a) protein knockdown. Although the *LPA* sgRNA has only one exact match in the *LPA* transgene, which is at the on-target site in KIV_1, many potential off-target sites within the *LPA* gene on other KIV domains share high sequence homology. We hypothesized that efficient simultaneous cleavages at on-target and off-target sites within the *LPA* transgene could cause large deletions, leaving only a small amount of KIV_1 unrearranged. To test this hypothesis, we checked the frequency of KIV_1 unrearranged on-target sites by ddPCR (Figures 3A and S7). The unrearranged KIV_1 region normalized by *LPA* construct copy number was 36.8% ± 4.2% and 36.7% ± 1.8% in edited male and female animals, respectively. This is compared with 98.8% ± 3.7% and 101.1% ± 2.9% in control male and female animals, respectively, indicating frequent rearrangement at on-target double-strand breaks (DSBs) (Figure 3A).

For a comprehensive understanding of the chromosomal aberrations within the *LPA* transgene, we performed chromosomal aberrations analysis by single targeted linker-mediated PCR sequencing (CAST-seq) in control (Figures 3B and 3C) and edited (Figures 3D and 3E) female mice.[28] A bait primer binding upstream (PAM-proximal side) of the on-target site on KIV_1 was used to identify chromosomal aberrations derived from on-target and off-target activities of CRISPR by mapping the chromosomal sequences fused to PAM-proximal arm of the on-target DSB. Alignment of CAST-seq reads to the LPA transgene revealed the frequent chromosomal deletions and inversions between CRISPR-induced on- and off-target DSBs within the *LPA* transgene and small indels at the breakpoint junctions in edited mouse, all contributing to *LPA* disruption (Figures 3D and 3E). The most retrieved CAST-seq reads included large deletions between the on-target site in KIV_1 and off-target sites in $KIV_{2(1\text{-}5)}$ (OT10, OT12, OT13, and OT18 in Table S1) domains. The second most frequent event involved large deletion and inversion with an off-target site on KIV_5 (OT2 in Table S1). CAST-seq also revealed large deletions between the on-target site and off-target sites in KIV_{10} (OT6), KIV_6 (OT4), KIV_9 (OT3), and KIV_3 (OT9) (Figures 3D and 3E; Table S1).

To validate the CAST-seq-identified local chromosomal rearrangements between the on-target site in KIV_1 and potential off-target site on KIV_5, we used a forward primer specific to the on-target site PAM upstream region and a reverse primer specific to KIV_5 off-target site PAM downstream region (Figures S8A and S8B). We detected a full-length amplicon (3,211 bp) in control animals and truncated amplicons in all AAV-CRISPR-treated animals, indicating frequent local chromosomal rearrangement occurring between KIV_1 and KIV_5 (Figure 3E). Truncated amplicons of 1,174 and 490 bp are the product of a 2.7-kb large deletion created by on- and off-target DSBs re-joining (Figures S8A and S8B) or a 2-kb large deletion between two off-target DSBs at $KIV_{2(1)}$ and KIV_4 (OT15 in Table S1; Figures S8C and S8D).

To further investigate these complex rearrangements, we performed PCR amplification of full-length LPA cDNA (5,219 bp) followed by Pacific Biosciences (PacBio) Single-Molecule, Real-Time (SMRT) sequencing.[29] High-fidelity (HiFi) reads from AAV-GFP and AAV-CRISPR samples were aligned to the 13K LPA cDNA reference sequence (Figures S9A and S9B), and diverse insertions and deletions were analyzed (Figure S9C). In the AAV-GFP control sample, 70% of reads had expected full-length 13K cDNA sequence and remaining reads had unexpected homologous insertions or deletions, possibly due to the homologous recombination in the germline of founder mice[30] or limitation of long-range PCR across the repetitive regions (Figures S9A, S9C, and S10).[31] In the AAV-CRISPR-treated sample, 22% of reads had the expected full-length 13K cDNA sequence and an impressive 72% of reads showed large deletions ranging from 1,710 to 4,079 bp in increments of 342 bp, corresponding to the individual KIV domain size (Figures S9B and S9C). Deletions in the AAV-CRISPR sample closely matched CRISPR DSBs at on- and off-target sites, further confirming the ligation of these sequences identified through CAST-seq (Figure S9B). We also observed alleles containing multiple deletions at on- and off-target cut sites and between separate off-target sites (Figure S9B). Collectively, these data indicate that off-target cutting within the *LPA* transgene causes chromosomal deletions and rearrangements, explaining the robust degree of apo(a) knockdown despite a lack of small indels at the on-target site.

Identification of off-target sites within the context of the human genome

To investigate the efficiency and safety of our *LPA* targeting gene therapy in the context of the human genome, we performed in depth on- and off-target analysis in HEK293T cells (Figure 4). We electroporated px601 plasmid encoding *LPA* sgRNA/SaCas9 or *LPA* sgRNA/SaCas9 ribonucleoprotein (RNP) into HEK293T at 1–4 μg and 30–90 pmol dose, respectively. T7E1 showed minimal indel activity and no evidence of editing by ICE analysis in HEK293T (data not shown), similar to what we observed in transgenic mice (Figure S6). The ddPCR copy number assay was optimized for the human *LPA* gene, and we quantified the rearrangement at the on-target site in exon 2. We detected up to 92% on-target rearrangement with a dose of 4 μg of plasmid (Figure 4A) and up to 90% on-target rearrangement with a dose of 90 pmol RNP demonstrating highly active *LPA* sgRNA in HEK293T cells (Figure 4C). Next, we performed genome-wide unbiased identification of DSBs enabled by sequencing (GUIDE-seq) in HEK293T cells treated with electroporation of CRISPR-Cas9 as plasmid (Figure 4B) or RNP (Figure 4D), along with a blunt double-stranded oligodeoxynucleotides (dsODN). GUIDE-seq relies on capture of a dsODN into DSBs to identify CRISPR-induced off-target activities in the genome globally by sequencing.[32,33] Positive linear correlations between GUIDE-seq read counts and indel mutation frequencies have been reported.[33]

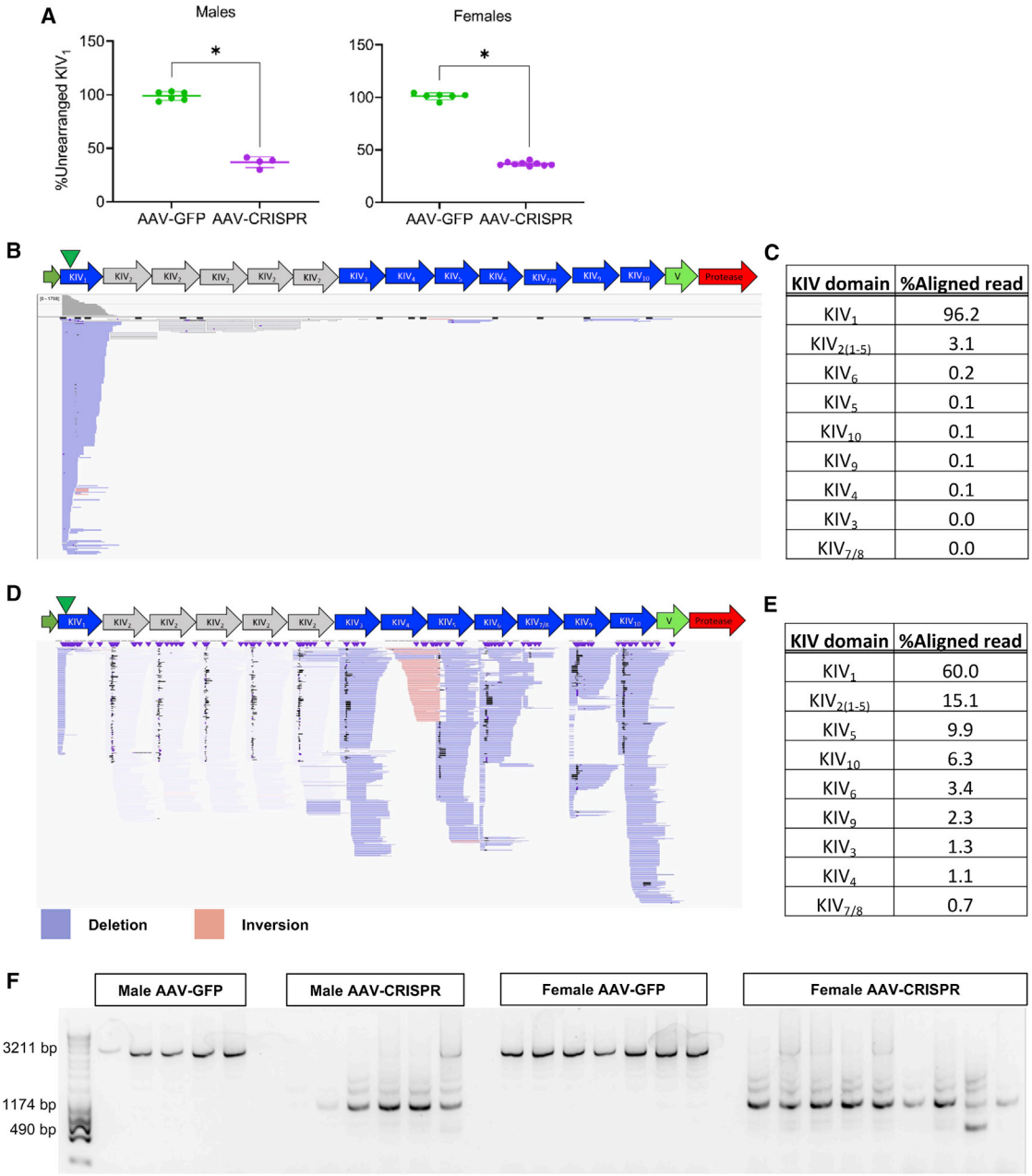


Figure 3. Chromosomal rearrangements in *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ mouse livers injected with AAV-CRISPR

(A) On-target unrearranged KIV1 normalized by *LPA* construct copy number, indicating frequent rearrangement at on-target DSB in AAV-treated mice. Statistical analysis was done using Welch's t test with *$p < 0.05$. Values are mean ± SD. (B–E). For CAST-seq, a bait primer binding upstream of the on-target site on KIV1 was used to identify chromosomal aberrations derived from on-target and off-target activities of CRISPR. CAST-seq reads aligned to the 13K *LPA* are represented by bars in Integrated Genomics Viewer (IGV) plots. The green triangle marks the position of the on-target site in KIV_1. (B and C) In AAV-GFP sample, 96% of reads align to KIV_1 with a minimal background at other KIV domains indicating intact KIV_1 without chromosomal rearrangement. (D and E) Frequent chromosomal deletions and inversions between CRISPR-induced on- and off-target DSBs within the *LPA* transgene and small indels at the breakpoint junctions in AAV-CRISPR sample, all contributing to *LPA* disruption. The percentage of aligned CAST-seq read at each KIV domain in descending order. Fifteen percent of CAST-seq reads aligned at KIV_2 domains were displayed in gray due to low alignment scores as those reads are equally likely to map to five KIV_2 domains; 9.92% of reads aligned at KIV_5 are shown in blue or red, indicating deletion and inversion of sequences between KIV_1 and KIV_5, respectively. CAST-seq also revealed deletions between the on-target site and off-target sites in other KIV domains. (F) Agarose gel image showing a full-length PCR amplicon between KIV_1 and KIV_5 (3,211 bp) in AAV-GFP-treated mice, and truncated amplicons in all AAV-CRISPR-treated mice.

In the plasmid-treated sample, GUIDE-seq identified 25 off-target sites, which were labeled according to the descending order of read count (OT1–OT25) (Figure 4B and Table S1). Of 25 off-target sites, 11 sites with one to four mismatches (MMs) to the target site are at LPA exons corresponding to KIV_{2-10}, five off-target sites with two to four MMs are at pseudogenes nearby LPA, and nine off-target sites with three to seven MMs and low read counts are at other loci. Due to highly homologous KIV domains, some of the off-target sites (e.g., OT4 and OT5; OT7 and OT8; OT9 and OT10; OT12, OT13, and OT18) have the same target and PAM sequences but occur at different genomic locations. The off-target site with the highest read count, OT1, has one MM on PAM and three PAM-distal MMs and is located at *LOC107986665* pseudogene, 256 kbp away from *LPA* on-target site (Table S1; Figure S11). In the RNP-treated sample, a total of 13 off-target sites were identified, with 12 sites also found in the plasmid-treated sample (Figures 4D and 4E). GUIDE-seq performed on the HEK293T cells receiving CRISPR-Cas9 delivered as a plasmid and RNP identified the overlapping off-target sites at or near LPA with high abundance, but the plasmid-treated sample generated a higher number of identified off-target sites as well as higher ratio of GUIDE-seq read count normalized by that of on-target sites (Figure 4E).

Analysis of chromosomal rearrangements as a result of *LPA* gene editing within the context of the human genome

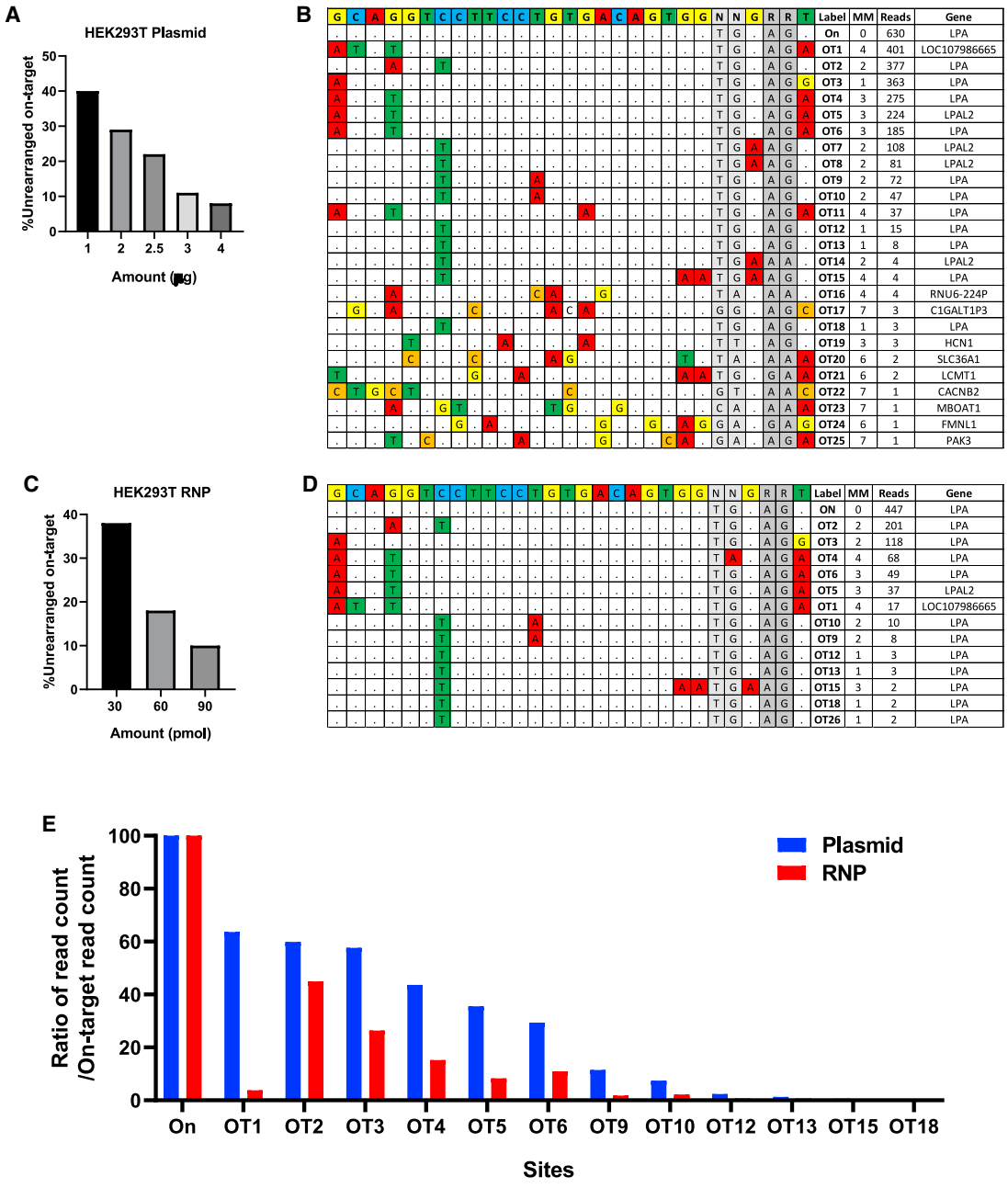
Genomic DNA from plasmid- (Figures 5A and 5B) and RNP (Figures 5C and 5D)-treated HEK293T cells were subjected to CAST-seq to identify chromosomal rearrangements involving the on-target DSB in the human genome (hg38). CAST-seq performed using plasmid-treated HEK293T cells identified 14 off-target rearrangements, of which eight are located at *LPA*, five at nearby pseudogenes, and one at *PLG* (Figure 5A). CAST-seq performed using RNP-treated HEK293T cells identified nine off-target rearrangements, of which eight were found in the plasmid-treated sample (Figure 5C). CAST-seq showed frequent rearrangements between on- and off-target sites within the *LPA*, particularly at OT2 on exon 20 and OT3 on exon 28 (Figures 5A and 5C; Table S1). The on- and off-target rearrangement identified by CAST-seq in plasmid- (Figure 5B) and RNP (Figure 5D)-treated samples was in good agreement with off-target DSBs identified by GUIDE-seq, demonstrating efficient chromosomal rearrangement between on- and off-target DSBs. CAST-seq-identified chromosomal rearrangement events were validated by endpoint PCR using the primers designed to amplify across the expected large deletion junction between the PAM upstream side of on-target DSB and the off-target DSBs at *LPA* (OT2, OT3, OT4, OT6, OT12, OT13, OT18, and OT26) or pseudogene (OT1) (Figure 5E). Next, we adapted a KIV_2 copy number variant (CNV) genotyping qPCR assay[34] for ddPCR and quantified the KIV_2 repeat number in unmodified and edited HEK293T. The primers and probe used for the assay were specific for KIV_2 and KIV_3 domains. We found that HEK293T contains ~18 c/dg of KIV_2 and KIV_3 domains, which decreased to 5 c/dg with 4 μg of plasmid delivery and was also reduced in a dose-dependent manner with RNP treatment (Figure 5F). Since *LPA* harbors intragenic multiallelic CNV and the number of KIV_2 copies varies from one to >40, these results indicate that the number of CRISPR-induced DSBs and extent of chromosomal arrangement would be influenced by KIV_2 CNV polymorphism in the individual genome.

DISCUSSION

Despite the causal role of Lp(a) in ASCVD and CAVD disease processes, there is currently no US Food and Drug Administration (FDA)-approved therapy that specifically targets Lp(a). Efforts to inhibit Lp(a) production by the liver through antisense and RNA interference approaches are promising, but it remains to be seen if these drugs will decrease cardiovascular events and associated mortality. Here, we present the first evidence that CRISPR-Cas9 can be used to permanently disrupt the *LPA* gene *in vivo*, rapidly and efficiently lowering plasma apo(a) levels.

We developed a new transgenic mouse model that expresses physiologically relevant levels of apo(a) in the circulation. Although intended to be liver specific, expression of *LPA* mRNA was also detected in the brain, which is likely a consequence of the *APOE* regulatory elements contained in the pLIV promoter. It is unlikely that brain-derived apo(a) could be secreted and enter the systemic circulation, as apoB is not expressed in this tissue, and apoE-containing lipoproteins constitute a pool that does not cross the blood-brain barrier.[35,36] This mouse model expresses a small isoform of *LPA* containing five KIV_2 domains, an important advantage for therapeutic testing and disease modeling as shorter isoforms are associated with more efficient secretion from the liver and higher Lp(a) levels in humans.[12] Although these studies were not performed in mice expressing human apoB-100, we found by size-exclusion chromatography of mouse plasma that apo(a) was localized in the VLDL and IDL/LDL fractions, with no detectable free apo(a) in plasma or on HDL. Note that KIV_7 and KIV_8 have been identified as key for non-covalent lysine-dependent interactions between apo(a) and apoB-100.[37,38] Most of KIV_8 is deleted in our transgene, but, since KIV_6 contains an identical lysine-binding site to KIV_7 and KIV_8, it appears that the KIV_6-KIV_7 sequence sufficiently compensates in non-covalent binding to mouse apoB-100. Future studies designed to assess the impact of *LPA* gene editing on atherosclerosis in mouse models should also include the human gene encoding apoB-100 to afford the production of *bona fide* covalent Lp(a) particles.

In humans, endogenous apo(a) is produced exclusively by the liver, making it an excellent target for therapeutic disruption with AAV-delivered CRISPR-Cas9. In order to achieve the greatest degree of knockout and specificity, we chose a sgRNA targeting exon 2 of the *LPA* transgene in the KIV_1 domain. AAV-CRISPR delivery to mice resulted in very efficient removal of apo(a) protein from the circulation, as demonstrated by western blotting and ELISA. This nearly complete loss of apo(a) (over 99% in male and over 96% in female mice) was stable for the 4-week observation period. Although sex differences in AAV gene delivery to the liver have been reported,[23] with males being more efficiently transduced, we obtained a comparable extent of apo(a) removal in female mice. This is likely due to the dose used,



Label	MM	Reads	Gene
On	0	630	LPA
OT1	4	401	LOC107986665
OT2	2	377	LPA
OT3	1	363	LPA
OT4	3	275	LPA
OT5	3	224	LPAL2
OT6	3	185	LPA
OT7	2	108	LPAL2
OT8	2	81	LPAL2
OT9	2	72	LPA
OT10	2	47	LPA
OT11	4	37	LPA
OT12	1	15	LPA
OT13	1	8	LPA
OT14	2	4	LPAL2
OT15	4	4	LPA
OT16	4	4	RNU6-224P
OT17	7	3	C1GALT1P3
OT18	1	3	LPA
OT19	3	3	HCN1
OT20	6	2	SLC36A1
OT21	6	2	LCMT1
OT22	7	1	CACNB2
OT23	7	1	MBOAT1
OT24	6	1	FMNL1
OT25	7	1	PAK3

Label	MM	Reads	Gene
ON	0	447	LPA
OT2	2	201	LPA
OT3	2	118	LPA
OT4	4	68	LPA
OT6	3	49	LPA
OT5	3	37	LPAL2
OT1	4	17	LOC107986665
OT10	2	10	LPA
OT9	2	8	LPA
OT12	1	3	LPA
OT13	1	3	LPA
OT15	3	2	LPA
OT18	1	2	LPA
OT26	1	2	LPA

Figure 4. GUIDE-seq off-target analysis in HEK293T cells treated with *LPA* sgRNA/SaCas9 plasmid and RNP

(A) HEK293T cells were electroporated using 1–4 μg of px601 plasmid encoding *LPA* sgRNA/*Sa*Cas9, and percentage of unrearranged on-target site was measured by ddPCR showing plasmid dose-dependent decrease in the unrearranged on-target site. (B) GUIDE-seq results from HEK293T cells treated by electroporation of 2.5 μg of plasmid along with 20 pmol of dsTag showing on-target site (On) and 25 off-target sites. The reference *LPA* on-target site is shown at the top (N, any nucleotide; R, purine). Nucleotide mismatches between on- and off-target sites are highlighted. Off-target sites were labeled according to the descending order of GUIDE-seq read count (see also Table S1). (C) HEK293T cells were electroporated using 30–90 pmol *LPA* sgRNA/*Sa*Cas9 as RNP. RNP dose-dependent decrease in the unrearranged on-target site is shown. (D) GUIDE-seq results from HEK293T cells electroporated with 90 pmol RNP and 20 pmol dsTag showing 13 identified off-target sites. (E) Ratio of GUIDE-seq read counts (on- or off-target read counts/on read count) for the sites identified by both plasmid and RNP-treated samples. Nearly all off-target sites identified by RNP (except OT26) were also found by plasmid, with a higher ratio of GUIDE-seq read counts indicating a higher degree of off-target mutagenesis in plasmid-treated sample. OT, off target.

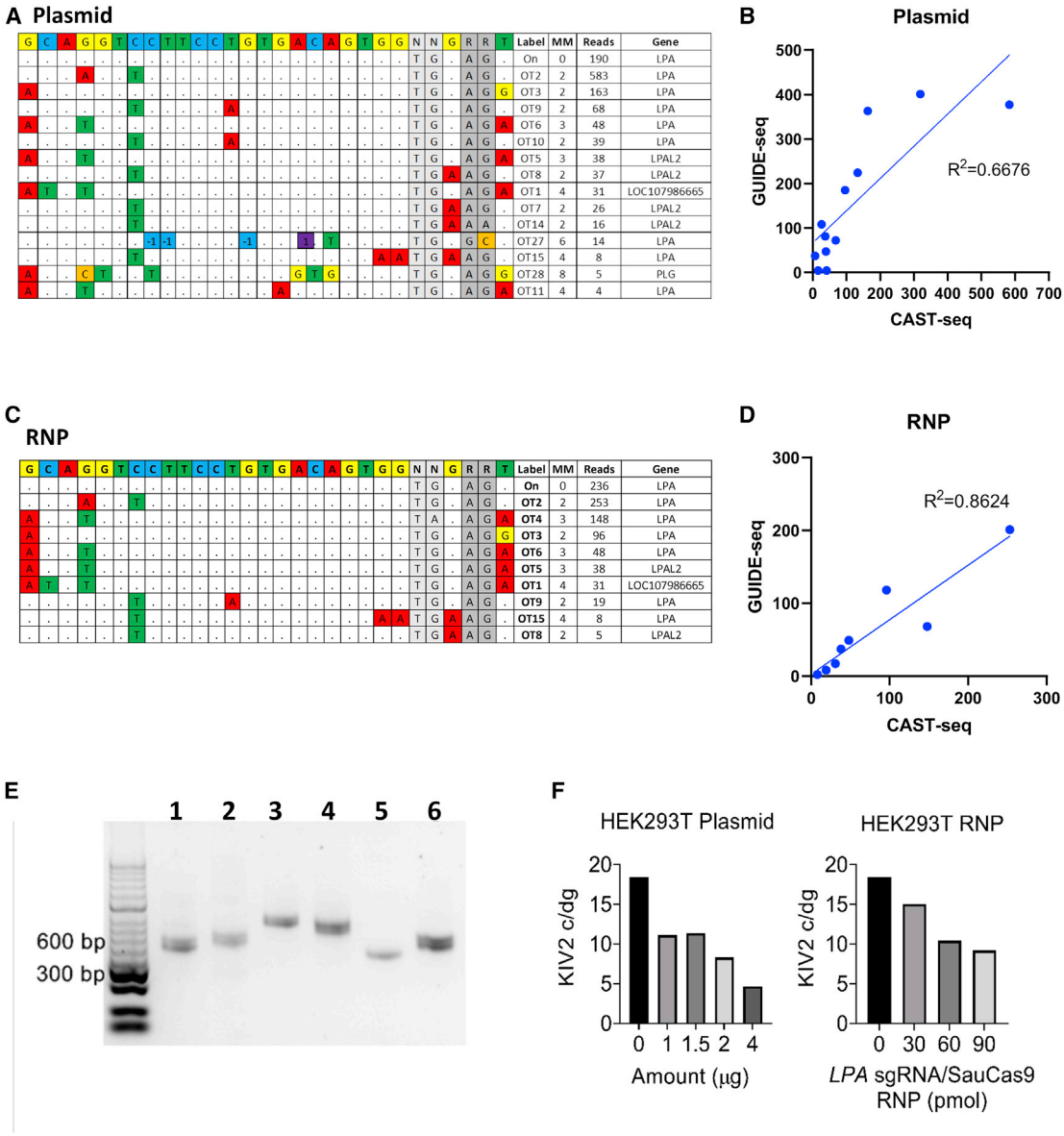


Figure 5. CAST-seq off-target analysis and on-target *LPA* rearrangements in HEK293T cells treated with *LPA* sgRNA/*Sa*Cas9 plasmid and RNP

(A) CAST-seq was performed on RNP and plasmid-treated HEK293T using the bait primer on the PAM upstream region of the on-target site to identify chromosomal rearrangements between on-target and off-target DSBs. CAST-seq identified on- and off-target chromosomal rearrangement in HEK293T cells electroporated with 2.5 μg of plasmid. The reference *LPA* on-target site is shown at the top. Mismatched nucleotides and indels (−1/1) are highlighted. (B) Correlation between GUIDE-seq read counts and CAST-seq read counts for the OT sites identified by both assays in plasmid-treated HEK293T. (C) CAST-seq identified on- and off-target chromosomal rearrangement in HEK293T cells electroporated with 90 pmol of RNP. (D) Correlation between GUIDE-seq read counts and CAST-seq read counts for the OT sites identified by both assays in RNP-treated HEK293T. (E) CAST-seq-identified chromosomal rearrangement events were validated by endpoint PCR using the primers designed to amplify across the expected large deletion junction between the PAM upstream side of on-target DSB and the off-target DSBs at selected OT sites. Lane 1, OT2 (404 bp); 2, OT3 (448 bp); 3, OT4 (579 bp); 4, OT6 (530 bp); 5, OT1 (376 bp); and 6, OT12, OT13, OT18, and OT26 (399 bp). PCR products in lane 6 could be from deletion between the on-target site and four different OT sites (OT12, OT13, OT18, and OT26) that share the same target sequence and PCR primer binding sequence. (F) ddPCR quantification of the KIV_2 CNV in untreated and plasmid- or RNP-treated HEK293T. The primers and probe used for the assay were specific for KIV_2 and KIV_3 domains. Untreated HEK293T contains ~18 c/dg of KIV_2 and KIV_3 domains, which decreased in a plasmid- and RNP dose-dependent manner.

which is approximately 5- to 10-fold higher than that required for complete hepatocyte transduction in male mice. Nonetheless, this is a clinically relevant dose of AAV8 (4×10^{13} GC/kg) that has been used in humans without major adverse events.[39] Evidence of safety was also supported by the lack of changes to overall body weight, ALT, AST, plasma cholesterol, as well as normal liver histology.

We were initially surprised to find no editing at the on-target site in the AAV-CRISPR mice by T7E1 or ICE analysis, given that we observed complete protein knockdown. Further analysis showed that many off-target edits occurred within the *LPA* transgene itself. Through ddPCR analysis, we determined that approximately 64% of *LPA* transgene alleles had a rearranged KIV_1-encoding domain. CAST-seq and PacBio long-read sequencing analysis revealed that off-target indels in other KIV domains in *LPA* and deletions between off-target sites also account for further loss of apo(a) expression. Considering that 20%–40% of the cells within the liver are non-parenchymal cells[40] that do not express *LPA*[41] but still contribute DNA to the pool, it stands to reason that we achieved editing in the vast majority of hepatocytes within the liver. Therefore, the primary effect of CRISPR editing in our model was segmental deletions of the *LPA* transgene mediated by on- and off-target cutting within similar KIV domains. AAV-CRISPR reduced plasma apo(a) to nearly undetectable levels mediated by cumulative consequences of *LPA* disruptive mutations.

Given the unanticipated editing events in the transgenic mouse model, we wanted to carefully assess this targeting strategy on the full *LPA* gene within the context of the human genome using human embryonic kidney (HEK293T) cells. Again, we saw no editing in the cells by T7E1 or ICE analysis, but a remarkable 92% on-target rearrangement by ddPCR. Using a combination of CAST-seq and GUIDE-seq analysis, we determined that there were a large number of chromosomal rearrangement events between on- and off-target DSBs within *LPA* itself. The difference in the off-target effects observed between using Cas9 expressing plasmid versus RNP is likely due to RNP's short active span.[42] A higher degree of off-target mutagenesis could be expected with AAV-CRISPR treatment due to robust expression of Cas9 at 4 weeks post injection (Figures 2G and 2H).

Human *LPA* harbors intragenic multiallelic CNV, and the number of KIV_2 copies varies from two to >40.[18] As such, the number of CRISPR-induced DSBs and extent of chromosomal arrangement would vary based on KIV_2 CNV polymorphism in the individual genome. The majority of identified off-target activity appears to be localized to *LPA* exons or pseudogenes near the *LPA* locus. GUIDE-seq did not detect off-target sites within the plasminogen (*PLG*) gene, nor were any potential sites in *PLG* identified *in silico* by COSMID[43] (data not shown). Nonetheless, CAST-seq in plasmid-treated samples identified deletion between *LPA* exon 2 and *PLG* intron 9, although with low read counts (Figure 5A), and we detected a 255-kb deletion between on-target and OT1 (Figure 5E), resulting in the removal of the entire *PLG* gene located between *LPA* and *LOC107986665* (Figure S11). The frequency of this deletion event is expected to be low in comparison with other edits. However, since plasminogen is a critical enzyme in fibrinolysis, *LPA* gene-editing approaches should be carefully assessed to avoid off-target editing or loss of *PLG*.

In summary, we report the first demonstration of successful *in vivo* editing of *LPA*, a major emerging target in CVD risk reduction. A one-time treatment with a gene-editing nuclease could provide permanent removal of Lp(a) for patients at the greatest risk. We also identified challenges in targeting such a polymorphic and repetitive gene with the CRISPR-Cas9 system. Despite the heterogeneity of editing events, we achieved very efficient apo(a) protein removal from the circulation, establishing valuable proof of concept for a new class of Lp(a) therapeutics.

MATERIALS AND METHODS

The data that support the findings of this study are available from the corresponding author upon request. The next-generation sequencing (NGS) data are accessible at NCBI's Sequence Read Archive available at: https://www.ncbi.nlm.nih.gov/sra.

Plasmid design and cloning

The 1162-pAAV-HLP-EmGFP-SpA plasmid, which encodes an emerald GFP driven by a liver-specific promoter was previously reported and is publicly available through Addgene (Watertown, MA).[44] An sgRNA targeting exon 2 of *LPA*, encoding the KIV_1 domain, was designed based on the presence of an NNGRRT PAM for *Staphylococcus aureus* Cas9 (SaCas9). Cloning of the sgRNA (GCAGGTCC TTCCTGTGACAGTGG) was accomplished by annealing oligonucleotides (Sigma-Aldrich, St. Louis, MO) and ligating into the BbsI site of 1313-pAAV-U6-SA-BbsI-MluI-gRNA-HLP-SACas9-HA-OLLAS-spA (Addgene, Watertown, MA)[44] to obtain 1718-pAAV-U6-SA-AMD-LPa-gRNA3-HLP-SACas9-HA-OLLAS-spA.[45] The total size of the recombinant AAV genome is 4,662 bp, including the inverted terminal repeats (ITRs). Complete plasmid sequences are available upon request.

AAV production

Recombinant AAV8 vectors were generated as previously described,[46] with several modifications.[24] Plasmids required for AAV packaging, adenoviral helper plasmid pAdDeltaF6 (PL-F- PVADF6), and AAV8 packaging plasmid pAAV2/8 (PL-T-PV0007) were obtained from the University of Pennsylvania Vector Core. Cell pellets were harvested and purified using a single cesium chloride density gradient centrifugation. Fractions containing AAV were pooled and then dialyzed against PBS using a 100-kDa Spectra-Por Float-A-Lyzer G2 dialysis device (Fisher Scientific, Hampton, NH) to remove the cesium chloride. Purified AAV (now referred to as AAV-GFP and AAV-CRISPR, respectively) were concentrated using a Sartorius Vivaspin Turbo 4 Ultrafiltration Unit and stored at −80°C until use. AAV titers were calculated after DNase digestion using qPCR relative to a standard curve of the transgene plasmid. Primers used for titer are included in Table S2.

Animals

The *LPA* transgene was constructed in the pLIV.7 vector.[47] The vector contains ~3 kb of the 5′ flanking region of the human *APOE* gene containing the promotor, exon 1, intron 1 and part of exon 2 of *APOE*, a 254-bp fragment from the 3′-flanking region of *APOE* containing the polyadenylation signal sequence, and a liver element consisting of a 1.7-kb fragment from the hepatic control region of the *APOE/APOC1* gene locus. An apo(a) cDNA encoding a 14-kringle form of apo(a) (i.e., with all KIV domains as well as the KV and

protease domains present and five copies of KIV_2) was excised from the corresponding pRK5 mammalian expression plasmid[48] by digestion with *Sal*I and inserted into pLIV.7 digested with *Xho*I. The 14K cDNA also harbors the rs3798220 SNP encoding an Ile-Met substitution in the protease-like domain.[49] A linear fragment encompassing the regulatory regions from pLIV.7 and the 14K cDNA were excised by digestion with *Sal*I and *Spe*I. Transgenic mice were generated by pronuclear injection of the fragment into a *C57BL/6* strain at the London Regional Transgenic and Gene Targeting Facility. Founder mice were screened for the transgene and presence of apo(a) in plasma by western blot (described below), and a strain of transgenic mice was established by subsequent breeding of a founder mouse with *C57BL/6* mice. The transgene was then bred into $Ldlr^{-/-}$ mice from The Jackson Laboratory (*B6.129S7-Ldlr*^{tm1Her}/J, stock number 002207, Bar Harbor, ME) to more closely mimic a human lipoprotein profile. The transgene was maintained in the hemizygous or null state ($LPA^{+/0}$ $Ldlr^{-/-}$ or $LPA^{0/0}$ $Ldlr^{-/-}$). The PCR-based amplification of full-length cDNA and PacBio single-molecule long-read sequencing confirmed the 13K cDNA with 5 KIV_2 domains and a KIV_7-KIV_8 hybrid within the $LPA^{+/0}LDLR^{-/-}$ mice. This is still a physiologically relevant form of apo(a) as it contains a small number of KIV_2 repeats, and it appears to non-covalently bind to murine apoB-100 likely because KIV_6 contains an identical lysine-binding site to KIV_7 and KIV_8. Mice were housed in a pathogen-free animal facility with a daylight cycle from 0700 to 1900 h. Animals were allowed free access to food and water. Mice were maintained on a standard laboratory diet (irradiated PicoLab Select Rodent 50 IF/6F, LabDiet product code 5V5R, St. Louis, MO). Both male and female mice were used in all experiments unless otherwise stated. Mice were injected intraperitoneally at 8 weeks of age with 1×10^{12} genome copies of AAV expressing either the control (AAV-GFP) or sgRNA and CRISPR-Cas9 (AAV-CRISPR). Fasting for lipid measurements was performed for 5 h, and plasma was obtained by retro-orbital bleeding with heparinized Natelson collecting tubes under the influence of isofluorane anesthesia. Plasma was collected weekly until 12 weeks of age, at which point animals were euthanized to harvest livers for analysis. All procedures were performed according to the regulations and with the prior approval of the Institutional Animal Care and Use Committees of the Baylor College of Medicine (protocol AN-6243) and The University of Western Ontario (protocols 2016-087 and 2020-104).

Genotyping

The *LPA* transgene was detected by PCR using the primers 5′-CC CCTGTGGTCCGGCAGTGCTACC-3′ and 5′-GGGATGGCAGACA AGCTGGC-3′, which amplify a 955-bp fragment encompassing KIV_{10}, KV, and part of the protease-like domain. The following β-globin control primer pair was used in the same reaction: 5′-CCAATCTGCTCACACAGGATAGAGAGGGCAGG-3′ and 5′-CCTTGAGGCTGCCAAGTGATTCAGGCCATCG-3′. The wild-type *Ldlr* allele was detected by PCR using the primers 5′-CCA TATGCATCCCCAGTCTT-3′ and 5′-GCGATGGATACACTCACT GC-3′, and the null *Ldlr* allele was detected using primer pair 5′-AAT CCATCTTGTTCAATGGCCGATC-3′ and 5′-CCATATGCATCC CCAGTCTT-3′. Cycling conditions were as follows: 5 min at 95°C followed by 35 cycles of 30 s at 95°C, 1 min at 60°C, and 1 min at 72°C, then 3 min at 72°C, and holding at 4°C.

RNA analysis

Mice were fasted for 6 h and then anesthetized using Avertin (0.015–0.017 mL/g body weight; intraperitoneal injection). Hearts and livers were perfused with 10 mL of PBS containing 100 U/mL heparin. The liver, kidney, spleen, heart, small intestine, brain, and quadriceps muscles were removed, weighed, snap-frozen in liquid nitrogen, and stored at −80°C. RNA was extracted from all tissues using TriZol reagent (Invitrogen, Waltham, MA) as per manufacturer's instructions, and 3 μg of RNA was reverse-transcribed to cDNA using the iScript Advanced cDNA Synthesis kit for RT-qPCR (Bio-Rad, Hercules, CA). Tissue *LPA* mRNA levels were determined by real-time qPCR using the SsoAdvanced Universal SYBR Green Supermix kit (Bio-Rad, Hercules, CA) and calculated using the $\Delta\Delta C_T$ method. The primers used were as follows: *LPA* 5′-GGCCTCCTTCTGAA CAAGAC-3′ and 5′-GAAGAGGATGCACAGAGAGGG-3′; *Gapdh* 5′-GCGACTTCAACAGCAACTCC-3′ and 5′-TAGCCGTATTCAT GGTCATACC-3′.

Western blotting

Mouse plasma (4 μL) was added to 16 μL of 4× SDS-PAGE sample buffer in the presence or absence of 4 μL of 1 M DTT. Samples were then boiled for 7 min, briefly centrifuged, and loaded onto 4%–20% polyacrylamide gradient mini-gels (Bio-Rad, Hercules, CA). After electrophoresis at 120 V for 90 min, proteins were electroblotted (100 V, 90 min) onto polyvinylidene fluoride (PVDF) membranes (Millipore IPFL00010). Membranes were then blocked in 5% (v/v) nonfat dry milk in 20 mM Tris pH 8.0, 150 mM NaCl, 0.1% (v/v) Tween 20 for 1 h, then incubated with a primary antibody against apo(a) in blocking buffer overnight. The primary antibody was either the monoclonal antibody A5[50] (1:5,000) or a rabbit monoclonal antibody (1:3,000) purchased from Abcam. After washing three times in 20 mM Tris pH 8.0, 150 mM NaCl, 0.1% (v/v) Tween 20, membranes were incubated with horseradish peroxidase-conjugated secondary antibody (sheep anti-mouse immunoglobulin [Ig] G or goat anti-rabbit IgG, as appropriate) (1:5,000) in blocking buffer for 1 h. After three washes, immunoreactive bands were detected using SuperSignal West Pico PLUS chemiluminescent substrate (Thermo Fisher, Waltham, MA) on a Chemi-Doc imager (Bio-Rad, Hercules, CA). Liver tissue was homogenized in ∼10 volumes of radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris pH 8.0, 1 mM EDTA, 1% Triton X-100, 0.1% sodium dodecyl sulfate, 0.5% sodium deoxycholate, 150 mM sodium chloride, and protease inhibitors [Roche 11836153001]) using a Bead Blaster 24 (Benchmark D2400). Protein concentrations were determined using BCA assay (Thermo-Pierce #23227). Liver lysates (38 μg for HA-Tag and 20 μg for GFP) were diluted in 4× LDS buffer (Life Technologies, Ref. NP0007) supplemented with 5% beta-mercaptoethanol and separated by SDS-PAGE using 4%–12% gradient gels (Life Technologies NP0322BOX) at 150 V for 1.5 h. Proteins were transferred to PVDF membranes for either 1.5 h (GFP blot) or 4 h (HA-Tag

blot), followed by blocking for 2 h at room temperature in a 2:1 solution of Intercept (PBS) Blocking Buffer (Li-Cor, 927-70001) and PBS with 0.05% Tween 20 (PBS-T). Primary antibodies to the HA-Tag (1:1,000, rabbit, C29F4, 3724, Cell Signaling), GFP (1:3,000, rabbit, A-11122, Thermo Fisher Scientific), and beta tubulin (1:500, mouse, University of Iowa Developmental Studies Hybridoma Bank E7) were diluted in 1% BSA in PBS-T and membranes were incubated overnight at 4 C. Then, blots were washed three times (10 min each) with PBS-T. Goat anti-rabbit 680-nm and anti-mouse 800-nm antibodies (1:15,000, 611-144-002-0.5, and 610-145-002- 0.5, Rockland) were incubated at room temperature for 1.5 h, washed three times (10 min each) with PBS-T, and imaged using an Odyssey Classic Imager (Li-Cor).

Plasma analysis and fast protein liquid chromatography

Plasma lipoprotein separation was performed by the Mouse Metabolic Core (MMC) at Baylor College of Medicine. Plasma pooled from four mice per group (250 μL total) was fractionated by fast protein liquid chromatography (FPLC) using two Superose-6 columns connected in series (Pharmacia FPLC System, Amersham Pharmacia Biotech, Amersham, United Kingdom) as previously described.[51–53] Total plasma cholesterol in lipoprotein fractions was measured using the Cholesterol E Kit (Wako Pure Chemical Industries, Mountain View, CA).

Histology

Mouse livers were fixed overnight in 10% formalin, washed in 70% ethanol for 48 h, and stored in 70% ethanol. Paraffin embedding, sectioning, and antibody staining were performed by the Texas Digestive Diseases Morphology core as previously described.[54] Slides were imaged at ×10 magnification on a Nikon Ci-L bright-field microscope at the Integrated Microscopy Core (Baylor College of Medicine). Ki67-and TUNEL-positive cell quantification was performed by manual count, and nuclei were quantified by ImageJ (https://imagej.nih.gov).

ALT and AST assays

Plasma ALT and AST were measured using the Teco Diagnostics ALT (SGPT) Liquid Reagent (A524-150) and Teco Diagnostics AST (SGOT) Liquid Reagent (A559-150), respectively. Plasma was diluted 1:10 and 10 μL of plasma dilution was mixed with 100 μL of reagent mix (five volumes of reagent 1 with one volume of reagent 2, prewarmed at 37 C). NADH absorbance at 340 nm was read at time 0 and every 2–4 min up to 28 min at 37°C using a Tecan Infinite M200 PRO plater reader. ALT and AST levels in plasma from mice with known liver damage (assay positive control) were around 200 and 300 U/L, respectively (data not shown).

Vector genome analysis

DNA was isolated from liver samples using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Serial dilution of a known standard was used as control. The qPCR was completed with 250 ng of DNA using primers against either emGFP 5′- GCATCGACTTC AAGGAGGAC-3′ and 5′- TGCACGCTGCCGTCCTCGATG-3′ or SaCas9 5′-GTACTACGAGGAAACCGGGAAC-3′ and 5′-GTTGTT GTAGAAGGAGGCGATAAAC-3′.

Apo(a) ELISA

Apo(a) was measured using the Mercodia Lp(a) ELISA kit (Mercodia, Winston-Salem, NC) with some modifications to the sample preparation. For sample preparation, 5 μL of sample and 5 μL of pre-treatment solution were mixed and incubated for 1 h at room temperature. Then 1 mL of sample buffer was added to the sample and pre-treatment solution after 1 h and mixed. After the test procedure was completed, the absorbance was then read at 450 nm on an Infinite M200 PRO (Tecan, Morrisville, NC). The U/L measurements were then converted to mg/dL by multiplying by 0.1254 according to the manufacturer's instructions. $LPA^{0/0}$ $Ldlr^{-/-}$ pooled controls were used to subtract background absorbance measurements from all samples in the time-course experiments.

Transgene copy number analysis by ddPCR

Fifty micrograms of liver tissue was perfused with PBS and genomic DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany). Concentrations were determined with the Qubit 4 Fluorometer and dsDNA broad-range assay kit (Thermo Fisher, Waltham, MA). Multiple EvaGreen PCR reactions were run simultaneously to quantify the copy number of target and reference loci. EvaGreen-based reaction mixes were prepared with 20 ng of genomic DNA template, 1× QX200 ddPCR EvaGreen Supermix (Bio-Rad, Hercules, CA), 100 nM target or reference primers, and 10 U of EcoRI-HF restriction enzyme in each 20 μL of reaction mix. The copy number of target loci was normalized by the copy number of reference loci to quantify the target loci copy number in mock and treated samples. For the KIV_2 CNV assay, multiplex probe-based reaction mixes were prepared with 20 ng of genomic DNA template, 1× ddPCR Supermix for Probes (Bio-Rad, Hercules, CA), 900 nM each KIV_2 and RPP30 primers, 250 nM each KIV_2 (FAM) and RPP30 (VIC) probes, and 10 U *Eco*RI-HF restriction enzyme in each 20 μL of reaction mix. The copy number of KIV_2 was normalized by RPP30 to quantify the KIV_2 CNV in mock and treated samples. PCR was performed according to the manufacturer's cycling protocol with optimized annealing temperature. Primer and probe sequences are provided in Table S3.

Adaptor ligation-mediated PCR

Genomic DNA from the $LPA^{+/0}$ $Ldlr^{-/-}$ mouse was fragmented with either the restriction enzyme EcoRI-HF (NEB, Ipswich, MA) or BamHI-HF (NEB, Ipswich, MA). Then, each digested genome was treated with the NEBNext Ultra II EndRepair/dA-Tailing Module (NEB, Ipswich, MA), followed by ligation of the CAST-seq adaptor to the fragments via the NEBNext Ultra II Ligation Module (NEB, Ipswich, MA), providing a bidirectional binding site on every fragment for a universal primer complementary to the adaptor sequence. Both a reverse primer complementary to the 5′ beginning of the *LPA* construct and a forward primer complementary to the 3′ end of the construct were designed such that endpoint PCR could be performed by pairing each one of these two primers with the universal primer

and using the ligated fragmented genomic DNA as template. PCR amplicons were gel extracted and analyzed by Sanger sequencing to identify the *LPA* transgene construct knockin location in the mouse genome (Figure S2A). A similar assay was performed using enzymatic fragmentation following the manufacturer's instructions (NEBNext Ultra II FS DNA Library Prep) and sequenced on NGS to identify tandem junctions (Figures S2B and S2C). The primer sequences used for the *LPA* transgene knockin identification are provided in Table S3.

CAST-seq

The CAST-seq assay was optimized for *LPA* sgRNA, including decoy primers and a bait primer for the PAM upstream side of DSB, following the protocol described previously.[28] Decoy primers that substantially blocked on-target amplification across the *LPA* target site were selected. Then 500 ng of genomic DNA from AAV-CRISPR- and AAV-GFP-treated male and female mice were subjected to enzymatic fragmentation, end repair, A-tailing, and CAST-seq adapter ligation according to the manufacturer's instructions (NEBNext Ultra II FS DNA Library Prep). The first PCR was performed using the primers complementary to the linker sequence (linker primer) and a bait primer. Decoy primers were introduced to ensure that only when the binding sites of the decoy primers were lost because of large deletion or translocation did second and third PCR steps amplify the full-length product for NGS. The second PCR utilized nested primers to reduce the non-specific amplification. The third PCR introduced the barcoded Illumina adaptor for sequencing (NEB, NEBNext Multiplex Oligos for Illumina). The libraries were loaded into an Illumina MiSeq Reagent Kit V2-500 cycle (Illumina, San Diego, CA) according to the manufacturer's instructions. The *LPA* transgene construct sequence was used to map NGS reads on Integrated Genome Viewer (IGV). For CAST-seq analysis in HEK293T, sequences were processed using the CAST-seq pipeline publicly available on GitHub (https://github.com/AG-Boerries/CAST-Seq).

GUIDE-seq

Samples were prepared for GUIDE-seq analysis, largely following the protocol described in Tsai et al. 2015 and Nobles et al. 2019.[32,33] Briefly, a 2.5-μg px601 plasmid expressing Cas9 and *LPA* sgRNA and 90-pmol SaCas9/*LPA* sgRNA RNP along with a 20-pmol iGUIDE dsODN tag were nucleofected into HEK293T using SF buffer and CM-130 program. Genomic DNA was extracted 4 days post electroporation using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). Then 500 ng of genomic DNA were subjected to enzymatic fragmentation, end repair, A-tailing, and Y-adapter ligation according to the manufacturer's instructions (NEBNext Ultra II FS DNA Library Prep). The first round of discovery PCR was done for sense and antisense strands individually using a primer specific to the dsODN tag (GSP1) and a P5 primer. The second round of discovery PCR was done individually for sense and antisense strands with nested primers for the dsODN (GSP2) and P5 primer. Samples were purified and concentrated using 0.7× beads (GE, Sera-Mag Select). Final library quantification and QC were done using Qubit dsDNA HS Assay Kit and KAPA Library Quantification Kit. Libraries were mixed in equal molar concentrations, loaded into a 300-cycle version 2 Illumina kit, and run on a MiSeq (Illumina, San Diego, CA). Sequences were processed using the iGUIDE pipeline publicly available on GitHub (https://github.com/cnobles/iGUIDE). GRCh38/hg38 was used as a reference genome.

T7E1 assay

PCR reactions were carried out with primers surrounding the expected cut site of the *LPA* gRNA using Q5 Hot Start High-Fidelity DNA Polymerase (NEB, Ipswich, MA). The purified PCR products of the unedited control and experimental samples were denatured and reannealed to produce heteroduplex DNA. The reannealed products were digested with T7E1 (NEB, Ipswich, MA) according to the manufacturer's protocol. The results were analyzed by agarose gel electrophoresis. The primer sequences used for the T7E1 assay are provided in Table S3.

ICE analysis

PCR reactions were carried out with primers surrounding the expected cutsite of the *LPA* gRNA using Q5 Hot Start High-Fidelity DNA Polymerase (NEB, Ipswich, MA). The purified PCR products of the unedited control and experimental samples were processed by Sanger sequencing. The resulting sequence trace files (.ab1) were then uploaded into the ICE Web tool (available at https://ice.synthego.com/#/) for indels quantification. The primer sequences used for the ICE assay are provided in Table S3.

Library preparation for PacBio sequencing

The long-range PCR was performed using genomic DNA (gDNA) from AAV-GFP- or AAV-CRISPR-treated mouse. PCR reaction contained 100 ng of gDNA, 200 nM of primers in 100-μL reaction (LongAmp Hot Start Taq 2X Master Mix, NEB). The PCR program consisted of initial denaturation (2 min at 94°C) and 30 cycles of denaturation (30 s at 94°C), annealing (30 s at 60°C), and extension (6 min at 65°C). Five-hundred nanograms of the long-range PCR amplicons were used for PacBio library preparation, which consists of DNA damage repair, end repair/A-tail, SMRTbell adaptor ligation (SMRTbell Express Template Prep Kit 2.0), nuclease treatment (SMRTbell Enzyme Clean Up Kit), and AMPure bead purification following the standard protocol. The SMRT-bell library was sequenced on a PacBio Sequel II 8M flowcell in circular consensus sequencing (CCS) mode following the standard protocol with 1-h pre-extension and 30-h collection time (Pacific Biosciences). The PacBio subreads were converted to HiFi reads; 48,824 and 56,878 HiFi reads with sequencing accuracy above 99.9% in AAV-GFP- and AAV-CRISPR-treated samples, respectively, were used for analysis.

Statistics

Animal numbers were estimated based on previous experience with each specific assay and expected effect size. No pre-randomization was performed, and researchers were not blind to genotype. Animals were sex- and age-matched for all experiments as detailed in the figure legends. Data were tested for normal distribution using the

Shapiro-Wilk normality tests. Differences between two groups were assessed with either Welch's t test for normally distributed data or Mann-Whitney test for non-normally distributed data as detailed in figure legends. All data are shown as the mean ± standard deviation (SD). Comparisons involving two groups was done using Welch's t test with *p < 0.05. For comparisons involving three or more groups, an ANOVA with *post hoc* Tukey test was applied and *p < 0.05. GraphPad Prism 7 was used for statistical analyses.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon request.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2022.10.009.

ACKNOWLEDGMENTS

We would like to thank Dr. Pradip Kumar Saha and the Mouse Metabolic Core at Baylor College of Medicine for performing the fast performance liquid chromatography. This work was supported by American Heart Association fellowships (19PRE34380467 to A.M.D. and 19POST34430092 to M.D.G.), National Institutes of Health (R01HL152314 and UG3HL151545 to G.B.; U42OD026645, R01HL132840 and R01DK124477 to W.R.L.), NIH T32 training grant (T32HL07676 to A.M.D.), Canadian Institutes of Health Research project grant (PG-45375 to M.L.K. and M.B.B.), Heart and Stroke Foundation of Canada grant in aid (G-17-0018740 to M.L.K. and M.B.B.), and Natural Sciences and Engineering Research Council of Canada of Canada discovery grant (RGPIN-2015-05006 to M.L.K.). Imaging for this project was supported by the Integrated Microscopy Core at Baylor College of Medicine and the Center for Advanced Microscopy and Image Informatics (CAMII) with funding from NIH (DK56338, CA125123, ES030285) and CPRIT (RP150578, RP170719). This work was also supported by the Texas Digestive Diseases Morphology Core (P30DK56338).

AUTHOR CONTRIBUTIONS

A.M.D., S.H.P., J.M.A., G.B., M.B.B., M.L.K., and W.R.L. conceived the project and designed the studies. A.M.D. and S.H.P. performed and analyzed most of the experiments. A.E.H. produced the viral vectors. A.M.D., A.L., M.D.G., and M.C. conducted *in vivo* experiments. J.M.A., M.B.B., and M.L.K. generated the transgenic mice. J.M.A. performed RNA and western blot analysis. L.S. conducted experiments to assess gene-editing frequencies. A.Y. completed the adaptor ligation-mediated PCR. M.D.G. performed AST, ALT, and western blot experiments. A.M.D., S.H.P., M.B.B., M.L.K., and W.R.L. wrote the manuscript, which was revised and approved by all authors.

DECLARATION OF INTERESTS

M.L.K. has held a research grant from Pfizer; is a member of advisory boards for Amya Therapeutics and Novartis; has received speaker's honoraria/consulting fees from Amgen, Regeneron, and Eli Lilly; and holds/has held research contracts with Sanofi, Ionis, Eli Lilly, and Abcentra. M.B.B. has held a research contract with Ionis. S.M.M. has received consulting fees from Roche, Denka, and Novartis and research support from Amgen through Medpace.

REFERENCES

1. Berg, K. (1963). A new serum type in man- the Lp system. Acta Pathol. Microbiol. Scand. *59*, 369–382.
2. Berg, K., Dahlén, G., and Frick, M.H. (1974). Lp(a) lipoprotein and pre-β1-lipoprotein in patients with coronary heart disease. Clin. Genet. *6*, 230–235.
3. Djurovic, S., and Berg, K. (1997). Epidemiology of Lp(a) lipoprotein: its role in atherosclerotic/thrombotic disease. Clin. Genet. *52*, 281–292.
4. Emerging Risk Factors Collaboration, Erqou, S., Kaptoge, S., Perry, P.L., Di Angelantonio, E., Thompson, A., White, I.R., Marcovina, S.M., Collins, R., Thompson, S.G., and Danesh, J. (2009). Lipoprotein(a) concentration and the risk of coronary heart disease, Stroke, and nonvascular mortality the emerging risk factors collaboration * Europe PMC funders group. JAMA *302*, 412–423.
5. Erqou, S., Thompson, A., Di Angelantonio, E., Saleheen, D., Kaptoge, S., Marcovina, S., and Danesh, J. (2010). Apolipoprotein(a) isoforms and the risk of vascular disease. Systematic review of 40 studies involving 58, 000 participants. J. Am. Coll. Cardiol. *55*, 2160–2167.
6. Clarke, R., Peden, J.F., Hopewell, J.C., Kyriakou, T., Goel, A., Heath, S.C., Parish, S., Barlera, S., Franzosi, M.G., Rust, S., et al. (2009). Genetic variants associated with Lp(a) lipoprotein level and coronary disease. N. Engl. J. Med. *361*, 2518–2528.
7. Kamstrup, P.R., Tybjærg-hansen, A., Steffensen, R., and Nordestgaard, B.G. (2009). Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. JAMA *301*, 2331–2339.
8. Tsimikas, S., Fazio, S., Ferdinand, K.C., Ginsberg, H.N., Koschinsky, M.L., Marcovina, S.M., Moriarty, P.M., Rader, D.J., Remaley, A.T., Reyes-Soffer, G., et al. (2018). NHLBI working group recommendations to reduce lipoprotein(a)-mediated risk of cardiovascular disease and aortic stenosis. J. Am. Coll. Cardiol. *71*, 177–192.
9. Kamstrup, P.R., Tybjærg-Hansen, A., and Nordestgaard, B.G. (2014). Elevated lipoprotein(a) and risk of aortic valve stenosis in the general population. J. Am. Coll. Cardiol. *63*, 470–477.
10. Koschinsky, M.L., and Marcovina, S.M. (2004). Structure-function relationships in apolipoprotein(a): insights into lipoprotein(a) assembly and pathogenicity. Curr. Opin. Lipidol. *15*, 167–174.
11. McLean, J.W., Tomlinson, J.E., Kuang, W.J., Eaton, D.L., Chen, E.Y., Fless, G.M., Scanu, A.M., and Lawn, R.M. (1987). cDNA sequence of human apolipoprotein(a) is homologous to plasminogen. Nature *330*, 132–137.
12. Rader, D.J., Cain, W., Ikewaki, K., Talley, G., Zech, L.A., Usher, D., and Brewer, H.B. (1994). The inverse association of plasma lipoprotein(a) concentrations with apolipoprotein(a) isoform size is not due to differences in Lp(a) catabolism but to differences in production rate. J. Clin. Invest. *93*, 2758–2763.
13. Thanassoulis, G. (2019). Screening for high lipoprotein(a) the time is now. Circulation *139*, 1493–1496.
14. Nordestgaard, B.G., Langlois, M.R., Langsted, A., Chapman, M.J., Aakre, K.M., Baum, H., Borén, J., Bruckert, E., Catapano, A., Cobbaert, C., et al. (2020). Quantifying atherogenic lipoproteins for lipid-lowering strategies: consensus-based recommendations from EAS and EFLM. Atherosclerosis *294*, 46–61.
15. Reyes-Soffer, G., Ginsberg, H.N., Berglund, L., Duell, P.B., Heffron, S.P., Kamstrup, P.R., Lloyd-Jones, D.M., Marcovina, S.M., Yeang, C., and Koschinsky, M.L.; American Heart Association Council on Arteriosclerosis, Thrombosis and Vascular Biology; Council on Cardiovascular Radiology and Intervention; and Council on Peripheral Vascular Disease (2022). Lipoprotein(a): a genetically determined, causal, and prevalent risk factor for atherosclerotic cardiovascular disease: a scientific statement from the American heart association. Arterioscler. Thromb. Vasc. Biol. *42*, E48–E60.
16. Tsimikas, S., Karwatowska-Prokopczuk, E., Gouni-Berthold, I., Tardif, J.-C., Baum, S.J., Steinhagen-Thiessen, E., Shapiro, M.D., Stroes, E.S., Moriarty, P.M., Nordestgaard, B.G., et al. (2020). Lipoprotein(a) reduction in persons with cardiovascular disease. N. Engl. J. Med. *382*, 244–255.

17. Koren, M.J., Moriarty, P.M., Baum, S.J., Neutel, J., Hernandez-Illas, M., Weintraub, H.S., Florio, M., Kassahun, H., Melquist, S., Varrieur, T., et al. (2022). Preclinical development and phase 1 trial of a novel siRNA targeting lipoprotein(a). Nat. Med. *28*, 96–103.

18. Schmidt, K., Noureen, A., Kronenberg, F., and Utermann, G. (2016). Structure, function, and genetics of lipoprotein (a). J. Lipid Res. *57*, 1339–1359.

19. Emdin, C.A., Khera, A.V., Natarajan, P., Klarin, D., Won, H.-H., Peloso, G.M., Stitziel, N.O., Nomura, A., Zekavat, S.M., Bick, A.G., et al. (2016). Phenotypic characterization of genetically lowered human lipoprotein(a) levels. J. Am. Coll. Cardiol. *68*, 2761–2772.

20. Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. Nat. Commun. *9*, 1911.

21. Doudna, J.A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

22. Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186–191.

23. Jarrett, K.E., Lee, C.M., Yeh, Y.H., Hsu, R.H., Gupta, R., Zhang, M., Rodriguez, P.J., Lee, C.S., Gillard, B.K., Bissig, K.D., et al. (2017). Somatic genome editing with CRISPR/Cas9 generates and corrects a metabolic disease. Sci. Rep. *7*, 44624.

24. Jarrett, K.E., Lee, C., De Giorgi, M., Hurley, A., Gillard, B.K., Doerfler, A.M., Li, A., Pownall, H.J., Bao, G., and Lagor, W.R. (2018). Somatic editing of Ldlr with adeno-associated viral-CRISPR is an efficient tool for atherosclerosis research. Arterioscler. Thromb. Vasc. Biol. *38*, 1997–2006.

25. Chadwick, A.C., Evitt, N.H., Lv, W., and Musunuru, K. (2018). Reduced blood lipid levels with in vivo CRISPR-Cas9 base editing of ANGPTL3. Circulation *137*, 975–977.

26. Boffelli, D., Cheng, J.F., and Rubin, E.M. (2004). Convergent evolution in primates and an insectivore. Genomics *83*, 19–23.

27. Davidoff, A.M., Ng, C.Y.C., Zhou, J., Spence, Y., and Nathwani, A.C. (2003). Sex significantly influences transduction of murine liver by recombinant adeno-associated viral vectors through an androgen-dependent pathway. Blood *102*, 480–488.

28. Turchiano, G., Andrieux, G., Klermund, J., Blattner, G., Pennucci, V., el Gaz, M., Monaco, G., Poddar, S., Mussolino, C., Cornu, T.I., et al. (2021). Quantitative evaluation of chromosomal rearrangements in gene-edited human stem cells by CAST-Seq. Cell Stem Cell *28*, 1136–1147.e5.

29. Amarasinghe, S.L., Su, S., Dong, X., Zappia, L., Ritchie, M.E., and Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. Genome Biol. *21*, 30.

30. Würtele, H., Gusew, N., Lussier, R., and Chartrand, P. (2005). Characterization of in vivo recombination activities in the mouse embryo. Mol. Genet. Genomics. *273*, 252–263.

31. Hommelsheim, C.M., Frantzeskakis, L., Huang, M., and Ülker, B. (2014). PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. Sci. Rep. *4*, 5052.

32. Nobles, C.L., Reddy, S., Salas-Mckee, J., Liu, X., June, C.H., Melenhorst, J.J., Davis, M.M., Zhao, Y., and Bushman, F.D. (2019). IGUIDE: an improved pipeline for analyzing CRISPR cleavage specificity. Genome Biol. *20*, 14.

33. Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., et al. (2015). GUIDEseq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat. Biotechnol. *33*, 187–197.

34. Lanktree, M.B., Rajakumar, C., Brunt, J.H., Koschinsky, M.L., Connelly, P.W., and Hegele, R.A. (2009). Determination of lipoprotein(a) kringle repeat number from genomic DNA: copy number variation genotyping using qPCR. J. Lipid Res. *50*, 768–772.

35. Mahley, R.W. (2016). Central nervous system lipoproteins: ApoE and regulation of cholesterol metabolism. Arterioscler. Thromb. Vasc. Biol. *36*, 1305–1315.

36. Liu, M., Kuhel, D.G., Shen, L., Hui, D.Y., and Woods, S.C. (2012). Apolipoprotein E does not cross the blood-cerebrospinal fluid barrier, as revealed by an improved technique for sampling CSF from mice. Am. J. Physiol. Regul. Integr. Comp. Physiol. *303*, R903–R908.

37. Becker, L., McLeod, R.S., Marcovina, S.M., Yao, Z., and Koschinsky, M.L. (2001). Identification of a critical lysine residue in apolipoprotein B-100 that mediates noncovalent interaction with apolipoprotein(a). J. Biol. Chem. *276*, 36155–36162.

38. Becker, L., Cook, P.M., Wright, T.G., and Koschinsky, M.L. (2004). Quantitative evaluation of the contribution of weak lysine-binding sites present within apolipoprotein(a) kringle IV types 6-8 to lipoprotein(a) assembly. J. Biol. Chem. *279*, 2679–2688.

39. Pasi, K.J., Rangarajan, S., Mitchell, N., Lester, W., Symington, E., Madan, B., Laffan, M., Russell, C.B., Li, M., Pierce, G.F., et al. (2020). Multiyear follow-up of AAV5-hFVIII-SQ gene therapy for Hemophilia A. N. Engl. J. Med. *382*, 29–40.

40. Racanelli, V., and Rehermann, B. (2006). The liver as an immunological organ. Hepatology *43*, S54–S62.

41. Brancale, J., and Vilarinho, S. (2021). A single cell gene expression atlas of 28 human livers. J. Hepatol. *75*, 219–220.

42. Vakulskas, C.A., and Behlke, M.A. (2019). Evaluation and reduction of crispr off-target cleavage events. Nucleic Acid Ther. *29*, 167–174.

43. Cradick, T.J., Qiu, P., Lee, C.M., Fine, E.J., and Bao, G. (2014). COSMID: a web-based tool for identifying and validating CRISPR/Cas off-target sites. Mol. Ther. Nucleic Acids *3*, e214.

44. Li, A., Lee, C.M., Hurley, A.E., Jarrett, K.E., De Giorgi, M., Lu, W., Balderrama, K.S., Doerfler, A.M., Deshmukh, H., Ray, A., et al. (2019). A self-deleting AAV-CRISPR system for in vivo genome editing. Mol. Ther. Methods Clin. Dev. 111–122.

45. De Giorgi, M., Jarrett, K.E., de Aguiar Vallim, T.Q., and Lagor, W.R. (2022). In vivo gene editing in lipid and atherosclerosis research. In Atherosclerosis: Methods and Protocols, D. Ramji, ed. (Springer US), pp. 673–713.

46. Lagor, W.R., Johnston, J.C., Lock, M., Vandenberghe, L.H., and Rader, D.J. (2013). Adeno-associated viruses as liver-directed gene delivery vehicles: focus on lipoprotein metabolism. Methods Mol. Biol. *1027*, 273–307.

47. Fan, J., Wang, J., Bensadouni, A., Lauer, S.J., Dang, Q., Mahley, R.W., and Ii, J.M.T. (1994). Overexpression of hepatic lipase in transgenic rabbits leads to a marked reduction of plasma high density lipoproteins and intermediate density lipoproteins. Proc. Natl. Acad. Sci. USA, 8724–8728.

48. Gabel, B., Yao, Z., McLeod, R.S., Young, S.G., and Koschinsky, M.L. (1994). Carboxyl-terminal truncation of apolipoproteinB-100 inhibits lipoprotein(a) particle formation. FEBS Lett. *350*, 77–81.

49. Scipione, C.A., McAiney, J.T., Simard, D.J., Bazzi, Z.A., Gemin, M., Romagnuolo, R., Macrae, F.L., Ariëns, R.A., Hegele, R.A., Auld, J., et al. (2017). Characterization of the I4399M variant of apolipoprotein(a): implications for altered prothrombotic properties of lipoprotein(a). J. Thromb. Haemost. *15*, 1834–1844.

50. Marcovina, S.M., Albers, J.J., Gabel, B., Koschinsky, M.L., and Gaur, V.P. (1995). Effect of the number of apolipoprotein(a) kringle 4 domains on immunochemical measurements of lipoprotein(a). Clin. Chem. *41*, 246–255.

51. Chang, B.H.J., Liao, W., Li, L., Nakamuta, M., Mack, D., and Chan, L. (1999). Liver-specific inactivation of the abetalipoproteinemia gene completely abrogates very low density lipoprotein/low density lipoprotein production in a viable conditional knockout mouse. J. Biol. Chem. *274*, 6051–6055.

52. Son, S.H., Goo, Y.H., Choi, M., Saha, P.K., Oka, K., Chan, L.C.B., and Paul, A. (2016). Enhanced atheroprotection and lesion remodelling by targeting the foam cell and increasing plasma cholesterol acceptors. Cardiovasc. Res. *109*, 294–304.

53. Bissig-Choisat, B., Wang, L., Legras, X., Saha, P.K., Chen, L., Bell, P., Pankowicz, F.P., Hill, M.C., Barzi, M., Leyton, C.K., et al. (2015). Development and rescue of human familial hypercholesterolaemia in a xenograft mouse model. Nat. Commun. *6*, 7339.

54. De Giorgi, M., Jarrett, K.E., Burton, J.C., Doerfler, A.M., Hurley, A., Li, A., Hsu, R.H., Furgurson, M., Patel, K.R., Han, J., et al. (2020). Depletion of essential isoprenoids and ER stress induction following acute liver-specific deletion of HMG-CoA Reductase. J. Lipid Res. *61*, 1675–1686.

OMTM, Volume *27*

Supplemental information

LPA disruption with AAV-CRISPR potently lowers plasma apo(a) in transgenic mouse model: A proof-of-concept study

Alexandria M. Doerfler, So Hyun Park, Julia M. Assini, Amer Youssef, Lavanya Saxena, Adam B. Yaseen, Marco De Giorgi, Marcel Chuecos, Ayrea E. Hurley, Ang Li, Santica M. Marcovina, Gang Bao, Michael B. Boffa, Marlys L. Koschinsky, and William R. Lagor

Supplemental Information

Supplemental Figures and Legends

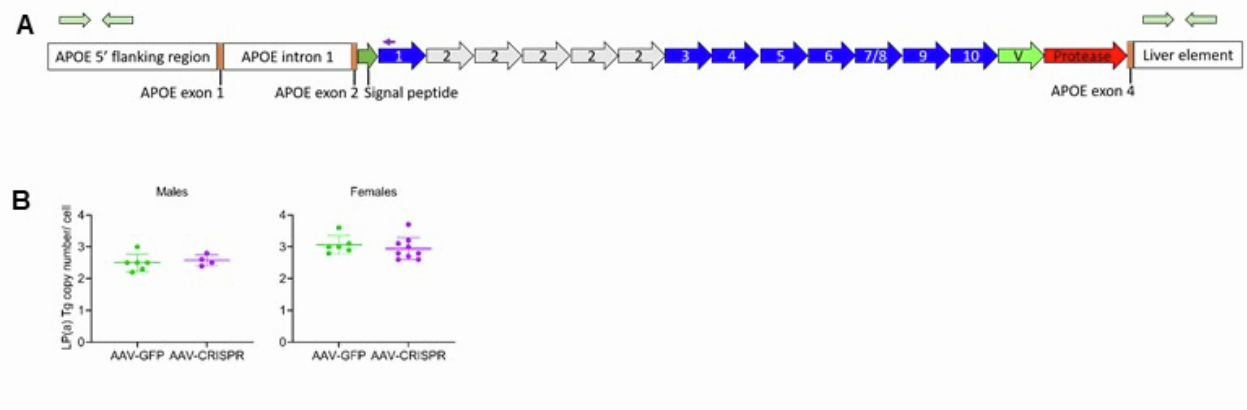


Figure S1. *LPA* transgene construct copy number analysis in *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ (See also Figure 1). A. Schematic of the *LPA* transgene construct and primers used for copy number analysis (indicated by green arrows, primer binding sites are not drawn to scale but indicate which elements on the construct they bind). B. ddPCR results showing an average of 3 *LPA* transgene copies/diploid genome (c/dg) in male and female *LPA*$^{+/0}$ Ldlr$^{-/-}$ mice treated with AAV-GFP or AAV-CRISPR. Statistical analysis was done using Welch's t-test with *p<0.05. Values are mean $\pm$ SD

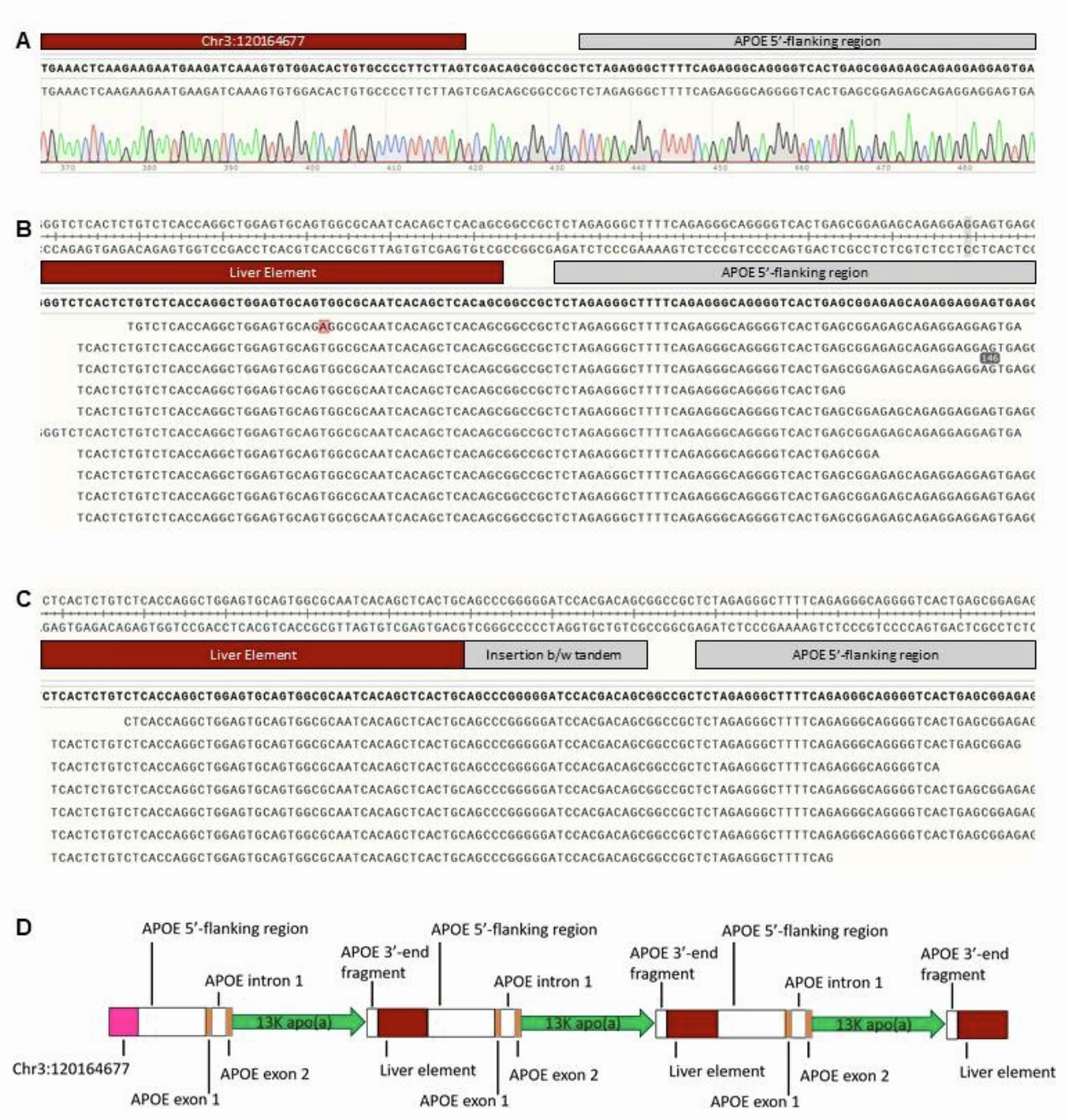


Figure S2. Sequencing results from adaptor ligation-mediated PCR method to identify *LPA* transgene knock-in location in *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ (see also Figure 1). A. Sanger sequencing results showing the junction of chr3:120164677 and the 5' end of the *LPA* construct. NGS results show the B. type 1 and C. type 2 tandem junction of the 3' and 5' end of the *LPA*

construct. D. Schematics of the integrated *LPA* transgene. *LPA*$^{+/0}$ Ldlr$^{-/-}$ mice have three tandem repeats of the *LPA* construct in direct orientations integrated at chr3:120164677.

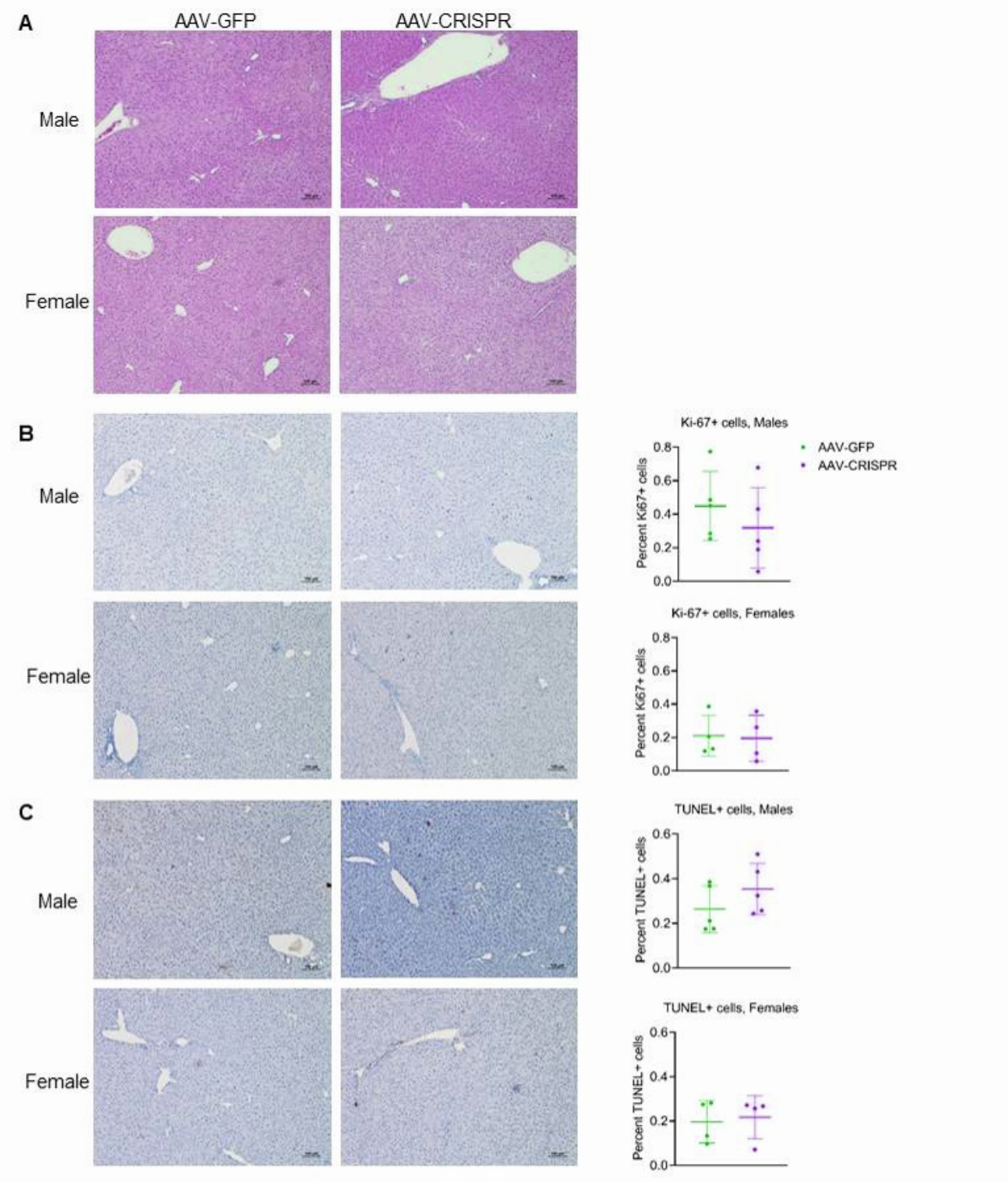


Figure S3. Histology analysis on AAV-GFP and AAV-CRISPR treated mice (See also Figure 2). A. Representative images for hematoxylin and eosin stained male and female mouse livers injected with either AAV-GFP or AAV-CRISPR. B. Representative images and

quantification of Ki67 stained male and female mouse livers. C. Representative images and quantification of TUNEL stained male and female mouse livers. Statistical analysis was done using Mann-Whitney test on non-normally distributed data with *p<0.05 (TUNEL+cells females). Statistical analysis was done using a Welch's *t*-test on normally distributed data with *p<0.05. (KI67+ cells males and females, and TUNEL+ cells males). Values are mean $\pm$ SD.

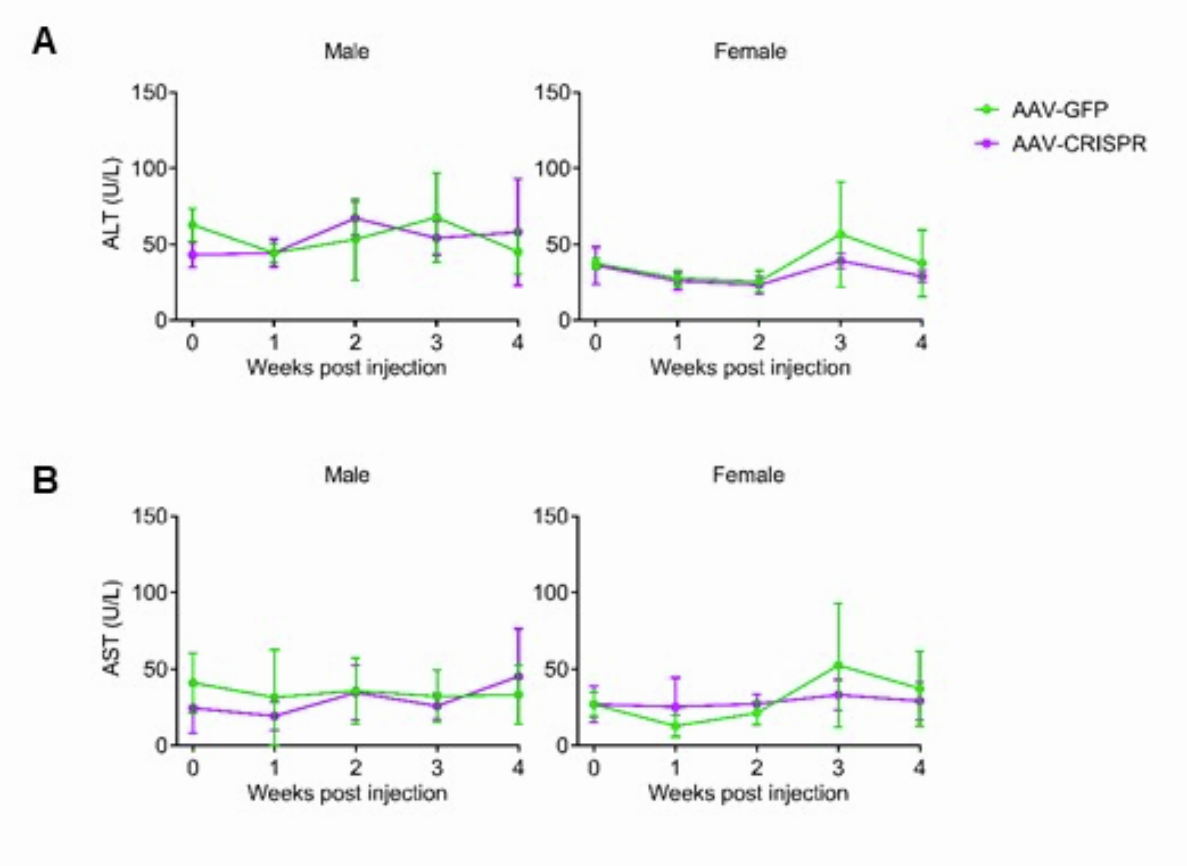


Figure S4. ALT and AST analyses (See also Figure 2). A. ALT analysis on plasma from mice treated with AAV-GFP or AAV-CRISPR. B. AST analysis on plasma from mice treated with AAV-GFP or AAV-CRISPR. Statistical analysis was done using a Welch's *t*-test at each time point with *p<0.05. Values are mean $\pm$ SD.

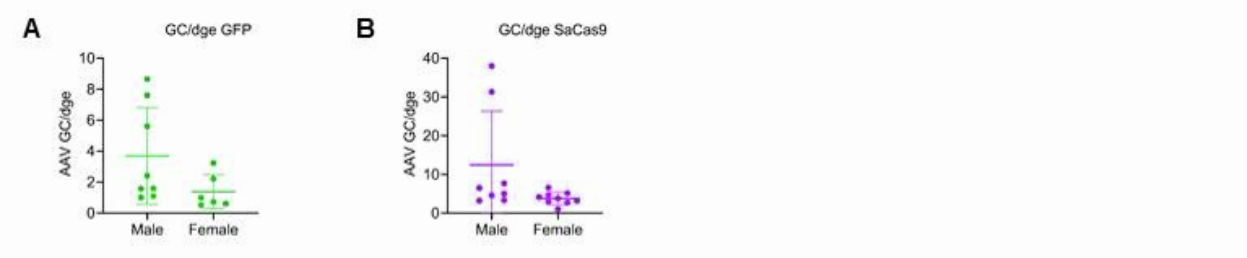


Figure S5. AAV vector genome copies per diploid genome analysis (See also Figure 2). qPCR for vector genomes in mice that received either A. GFP copy number in AAV-GFP treated mice or B. SaCas9 copy number in AAV-CRISPR treated mice. Statistical anlysis was done using Mann-Whitney test with *p<0.05. Values are mean $\pm$ SD.

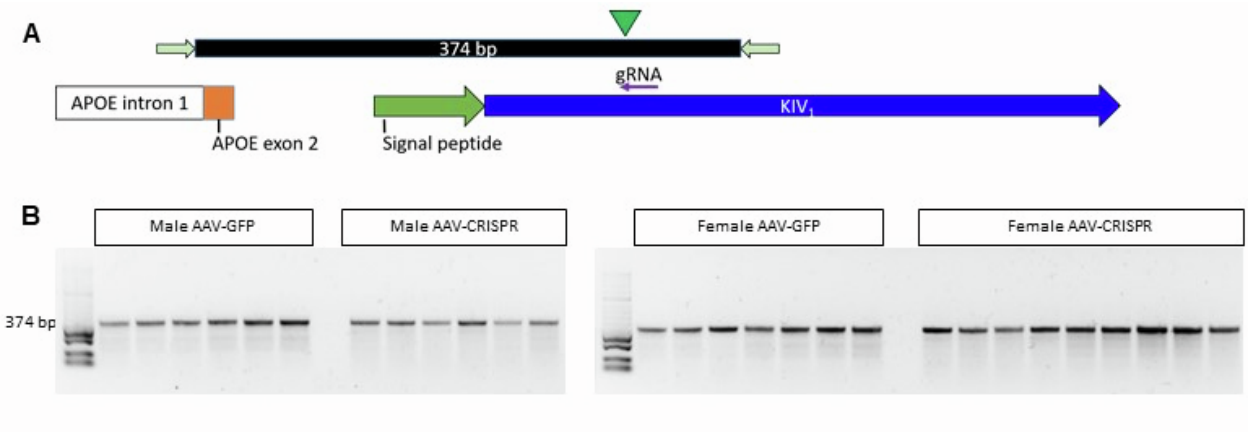


Figure S6. On-target insertion and deletion analysis of *LPA*$^{+/0}$ *Ldlr*$^{-/-}$ mouse livers injected with AAV-CRISPR (See also Figure 3). A. Schematics of KIV_1 specific primer design used for T7E1 and ICE analysis (indicated by green arrows, ICE analysis not shown). The on-target site is indicated by the green triangle. The purple arrow shows the 5' to 3' of gRNA specific for the KIV_1. B. No measurable cleavage activity was detected by T7E1 analysis.

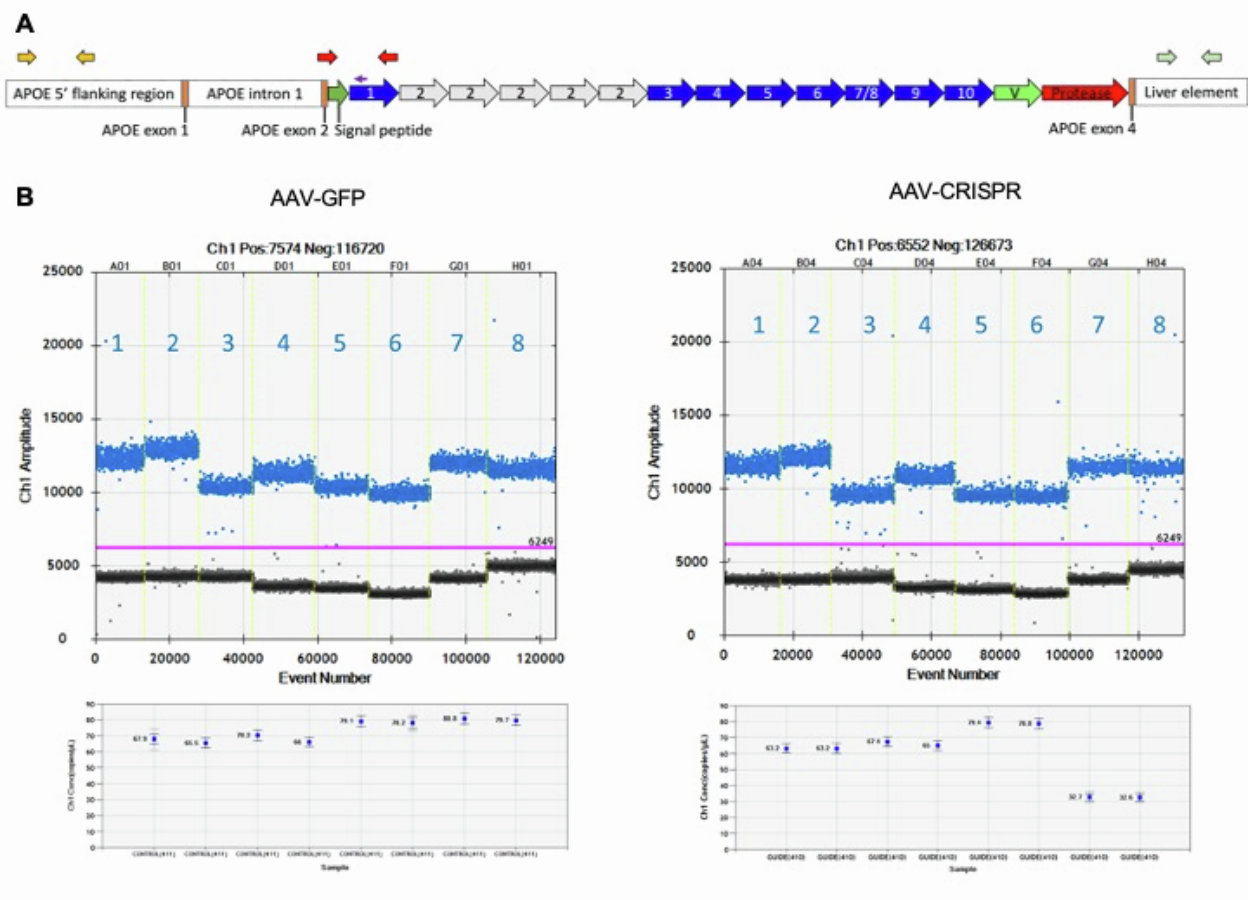


Figure S7. Representative Evagreen-based ddPCR copy number assay amplitude droplet separation and copy number quantification (See also Figure 3). A. Schematics of a KIV_1 specific primer pair used to quantify the frequency of un-rearranged on-target site on KIV_1 by ddPCR (indicated by red arrows). Yellow and green arrows indicate primer pairs used for copy number analysis at the APOE 5' flanking region or the liver element in the apo(a) transgene, respectively. (Primer binding sites are not drawn to scale but indicate which elements on the construct they bind). The purple arrow shows the 5' to 3' of gRNA specific for the KIV_1. The *TFRC* and *RPP30* genes were used as genomic reference sites for the ddPCR but primer binding sites are not shown. B. Representative ddPCR results for AAV- GFP and AAV-CRISPR treated mice. Lane 1: *TFRC* primer set 1, 2: *TFRC* primer set 2, 3: *RPP30* primer set 1, 4:

RPP30 primer set 2, 5: APOE 5’ end of *LPA* construct, 6: Liver element 3’ end of *LPA* construct, 7: *LPA* on-target in KIV_1 primer set 1 and 8: *LPA* on-target in KIV_1 primer set 2.

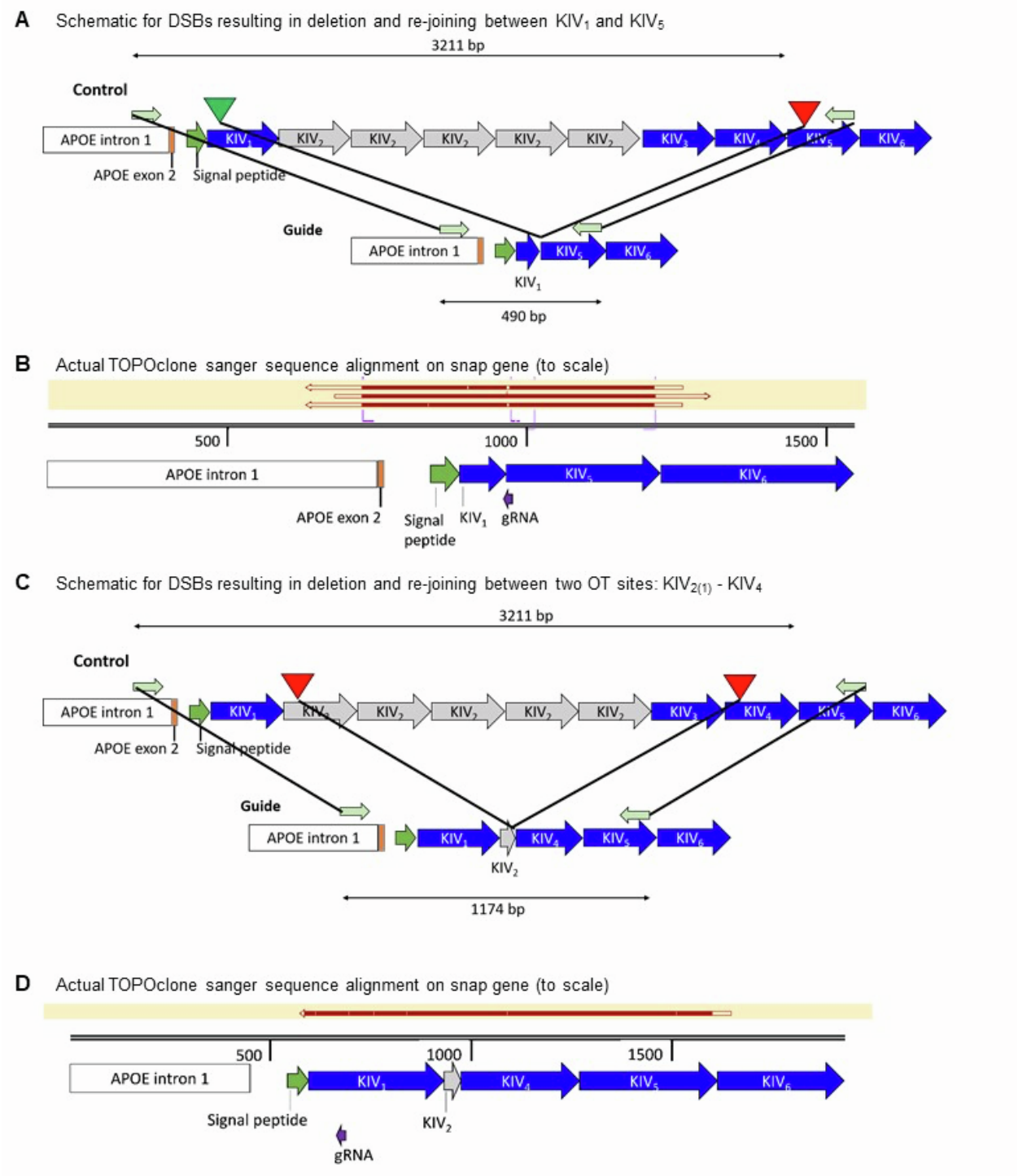


Figure S8. Schematics and aligned TOPO clone sequences for validation of CAST-seq identified large deletion between KIV_1 and KIV_5 (See also Figure 3). A. Schematic for TOPO clone sequencing of truncated amplicons from AAV-CRISPR treated mice showing the product

of 2.7kb large deletion created by DSBs re-joining or chromosomal rearrangements between KIV_1 and KIV_5 DSBs. B. TOPO clone sanger sequence alignment on SnapGene. C. Schematic for topo clone sequencing of truncated amplicons from AAV-CRISPR treated mice showing the product of 2kb large deletion created by DSBs re-joining or chromosomal rearrangements between two OT DSBs: $KIV_{2(1)}$ and KIV_4. D. TOPO clone sanger sequence alignment on SnapGene. On-target cut sites are indicated by green inverted triangles and OT cut sites are indicated by red inverted triangles. Specific primer pair design is indicated by light green arrows.

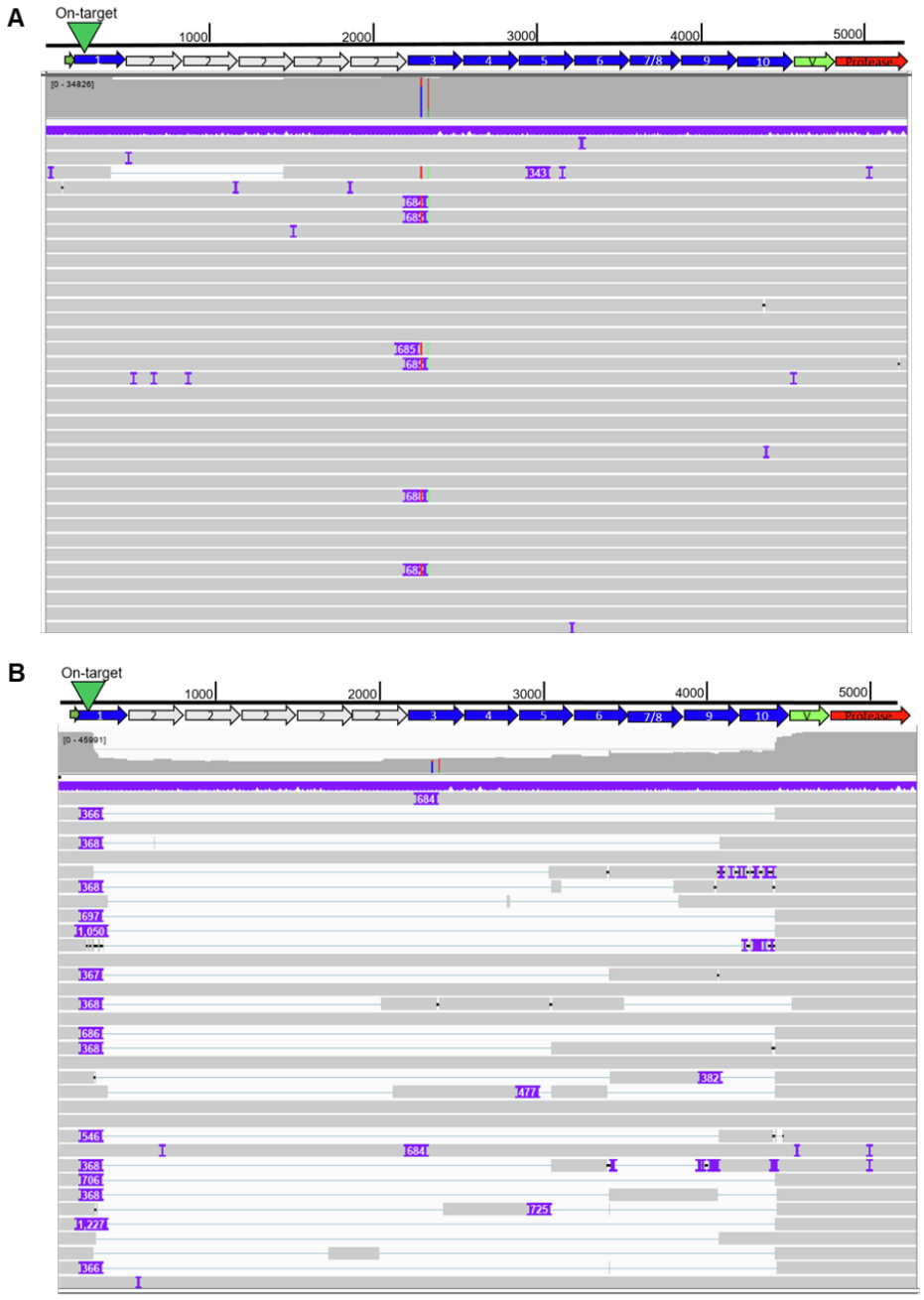

A
On-target
1000
2000
3000
4000
5000
Protease
B
On-target
1000
2000
3000
4000
5000
Protease
C
AAV-GFP
AAV-CRISPR
Read frequency%
80
60
40
20
0
-5000
-4800
-4600
-4400
-4200
-4000
-3800
-3600
-3400
-3200
-3000
-2800
-2600
-2400
-2200
-2000
-1800
-1600
-1400
-1200
-1000
-800
-600
-400
-200
0
200
400
600
800
1000
Large insertion or deletion sizes (bp)

Figure S9. PacBio SMRT sequencing analysis (See also Figure 3). PacBio SMRT sequencing generated HiFi reads alignment as viewed on IGV with either A. AAV-GFP or B.AAV-CRISPR treated samples. A. In AAV-GFP sample, 70% of reads were the expected full-length 13K cDNA, while 17% of reads had an unexpected longer 15K sequence with a 684 bp insertion at KIV_3 (**See Figure S10**). The remaining 12% of reads in the control sample were diverse homologous deletions of up to eight KIV domains most likely due to PCR artifacts. B. In AAV-CRISPR sample, 22% of reads had expected 13K cDNA sequences, while 6% of reads had 15K cDNA, and 72% of reads had deletions ranging from 1710 bp to 4079 bp. In contrast to the homologous deletions observed in AAV-GFP, deletions in the AAV-CRISPR sample are the products of CRISPR DSBs at ON and OT sites. We also observed alleles containing multiple deletions at ON-OT and between OTs. C. Histogram showing the frequency of reads corresponding to 13K, 15K, and deletions in AAV-GFP or AAV-CRISPR treated samples. Note that the read frequency does not represent the allele frequency due to template length bias during PCR amplification, but nonetheless provides information on a relative abundance and full-length sequences of LPA transgene alleles.

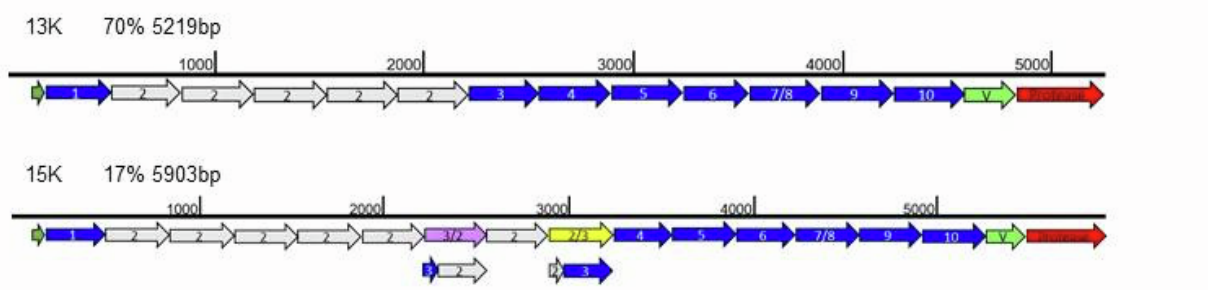


Figure S10. Schematic for transgene construct sequences for 13K or 15K (See also Figure 3). The longer 15K reads in the control sample contained a $KIV_{3/2}$ hybrid, $KIV_{2(6)}$, and $KIV_{2/3}$ hybrid contiguous sequence which could have resulted from homologous recombination in the germline of founder mice.[29] Another explanation for these unusual reads is a PCR artifact since amplification of highly similar sequences can produce undesired products such as recombination or chimera formation.[30] This remains to be validated using amplification-free long-read sequencing of the integrated LPA transgene.[1]

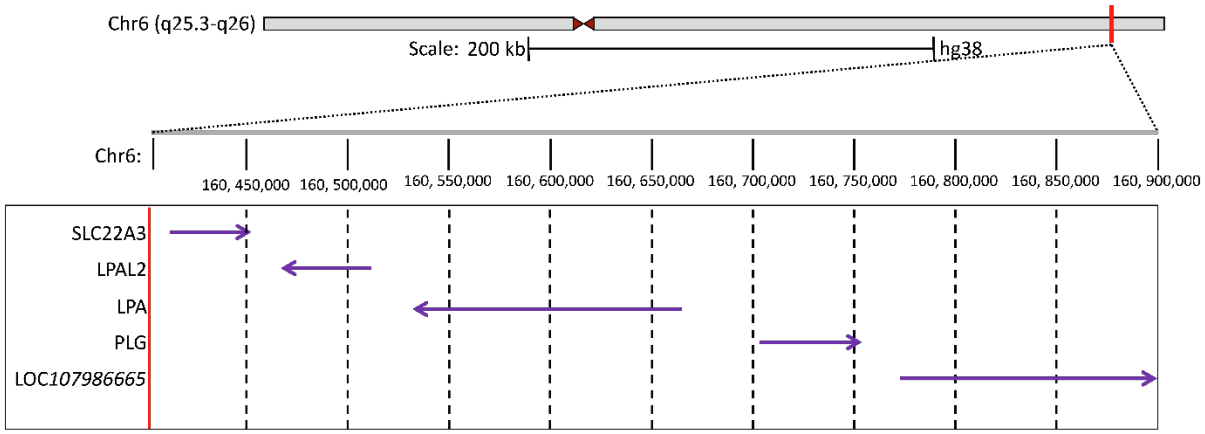


Figure S11. Schematic for the location of pseudogenes nearby *LPA* (See also Figure 4). Two of the GUIDE-seq identified off-target sites are located in *LPAL2* and *LOC107986665* pseudogenes, 173 kbp and 256 kbp away from *LPA* on-target, respectively. *PLG* is located between *LPA* and *LOC107986665.*

Supplemental Tables and Legends

Table S1. Off-target (OT) sites identified by GUIDE-seq in plasmid treated HEK293T (See also Figure 4). GUIDE-seq identified 25 off-target sites which were labeled according to the descending order of read abundance (OT1-OT25). For each off-target site, aligned sequences, location of edit site on Hg38, gene, GUIDE-seq identified read counts and number of mismatches (MMs) to the on-target site are shown. For the 11 off-target sites are at LPA, LPA exon number and corresponding to KIV domain are shown. Due to highly homologous KIV domains, some of the OT sites (ex. OT4 and OT5; OT7 and OT8; OT9 and OT10; OT12, OT13 and OT18) have the same target and PAM sequences but occur at different genomic locations.

Target	Aligned Sequence	Edit Site	Gene	LPA exon (KIV domain)	Distance from On (bp)	Reads	#MMs
On	.....TG.AG.	chr6:+:160650419	LPA	exon2 (KIV_1)	0	630	0
OT1	AT.T.....TG.AGA	chr6:-:160905840	LOC107986665		255421	401	4
OT2	...A..T.....TG.AG.	chr6:+:160599581	LPA	exon20 (KIV_5)	50838	377	2
OT3	A.....TG.AGG	chr6:+:160577217	LPA	exon28 (KIV_9)	73202	363	1
OT4	A..T.....TA.AGA	chr6:+:160594039	LPA	exon22 (KIV_6)	56380	275	4
OT5	A..T.....TG.AGA	chr6:+:160477210	LPAL2		173209	224	3
OT6	A..T.....TG.AGA	chr6:+:160556106	LPA	exon30 (KIV_{10})	94313	185	3
OT7	.....T.....TGAAG.	chr6:+:160500862	LPAL2		149557	108	2
OT8	.....T.....TGAAG.	chr6:+:160482741	LPAL2		167678	81	2
OT9	.....T.....A.....TG.AG.	chr6:+:160611643	LPA	exon16 (KIV_3)	38776	72	2
OT10	.....T.....A.....TG.AG.	chr6:+:160633834	LPA	exon8 (KIV_2)	16585	47	2
OT11	A..T.....A.....TG.AGA	chr6:+:160589634	LPA	exon24 (KIV_7)	60785	37	4
OT12	.....T.....TG.AG.	chr6:+:160617198	LPA	exon14 (KIV_2)	33221	15	1
OT13	.....T.....TG.AG.	chr6:+:160639381	LPA	exon6 (KIV_2)	11039	8	1
OT14	.....T.....TGAAA.	chr6:+:160491442	LPAL2		158977	4	2
OT15	.....T.....AATGAAG.	chr6:+:160605127	LPA	exon18 (KIV_4)	45292	4	4
OT16	...A.....CA..G.....TA.AA.	chr4:+:131259328	RNU6-224P			4	4
OT17	.G.A....C....ACA.....GG.AGC	chr8:+:62003251	C1GALT1P3			3	7
OT18	.....T.....TG.AG.	chr6:+:160644928	LPA	exon4 (KIV_2)	5491	3	1
OT19	....T.....A....A.....TT.AG.	chr5:+:45893210	HCN1			3	3
OT20	....C...C....AG.....T.GA.AAA	chr5:+:151448194	SLC36A1			2	6
OT21	T.....G..A.....AATG.GAA	chr16:+:25134130	LCMT1			2	6
OT22	CTGCT.....C.....GT.AAC	chr10:+:18249228	CACNB2			1	7
OT23	...A..GT.....TG..G.....CA.AAA	chr6:+:20125534	MBOAT1			1	7
OT24	.....G.A.....G..G.AGGAGAG	chr17:-:45243782	FMNL1			1	6
OT25	...T.C.....A....G...CA.GA.AGA	chrX:+:110905812	PAK3			1	7

Table S2. Primers used for AAV titer quantification.

Primer Name	Assay	Primer Sequence Sequence
WRL_0066	AAV titer SaCas9 Fw	GTACTACGAGGAAACCGGGAAC
WRL_0068	AAV titer SaCas9 Rv	GTTGTTGTAGAAGGAGGCGATAAAC
EYFP For2	AAV titer Cre/GFP Fw	GCATCGACTTCAAGGAGGAC
EGFP_Rev2	AAV titer Cre/GFP Rv	TGCACGCTGCCGTCCTCGATG

Table S3. List of primers used as well as corresponding figures and assays.

Mouse			
Primer name	**Figure**	**Assay**	**Primer Sequence**
KI_R1	FigS2	LPA transgene KI location	CACCAACAACTTCACGCTGG
KI_R2_nested			TCACTCCTCCTCTGCTCTCC
KI_F1			ACACTCAGAGCTTTGCAGCT
KI_F2_nested			CACTCTGTCTCACCAGGCTG
KIV1_F	Fig3A, FigS6, FigS7	ddPCR	TCACTGGCGGTTGATTGACA
KIV1_R			GCATTTGGGTAGTTTTCTGTGGT
KIV1_F2	Fig3F, FigS8	PCR and Sanger sequencing	TCACTGGCGGTTGATTGACA
KIV5_del_R			CCTCTGTGCTTGGATCTGGG
APOE_copy_F3	FigS1, FigS7	LPA transgene copy number ddPCR	GGAGAGCAGAGGAGGAGTGA
APOE_copy_R3			CCGACAGCTTCCAGACCTTT
Liver_copy_F4			CAACCCCTCAGTTCCCATCC
Liver_copy_R4			CCGAAATTCCAAGGGGTCGA
TFRC_new_F1			CTTCCCACCATGCTTGTTGC
TFRC_new_R1			CACTTCCGCTGCTGTACGAA
TFRC_new_F2			GCAGTGAGTTCTTCCCACCA
TFRC_new_R2			CGCTGCTGTACGAACCATTT
mRPP30_F1			CTACTGGCTTGGCTCTGTCC
mRPP30_R1			TAGGGACGTCAGGCAGATCA
mRPP30_F2			TCTGCAAGACGACCCAAGAC
mRPP30_F2			GGACAGAGCCAAGCCAGTAG
EYFP For2	Fig S5	AAV GC	GCATCGACTTCAAGGAGGAC
EGFP_Rev2			TGCACGCTGCCGTCCTCGATG
WRL_0066			GTACTACGAGGAAACCGGGAAC
WRL_0068			GTTGTTGTAGAAGGAGGCGATAAAC
Human			
Primer name	**Figure**	**Assay**	**Primer Sequence**
LPA_F1	Data not shown	T7E1, ICE	CTCAGGACCCAGGGTGTTTC
LPA_R1			CAGCTTGCAGTGCCAAATGT
LPA_long_F1	Fig4A,C	on-target unrearranged ddPCR	CACGTCTGTCTGCCTGCTAA
LPA_long_R1			AGTTAGCTTGACGCACACCT
LPA_long_F2			TTTTTGCACGTCTGTCTGCC
LPA_long_R2			GCTTGACGCACACCTTTTCT
CACNA1C_F			GAGAGTGCCTGGTCTTTACTGCAGG
CACNA1C_R			CTCTCCTATAGCTGCATGCACACCC
LPA_CAST_bait_F1	Fig5A,C	CAST-seq	TAGCAGCACCTGAGCAAAGC
LPA_CAST_bait_nest_F			CACCTGAGCAAAGCCATGTG
LPA_decoy_R			ACGCATTTGGGTAGTTTTCTGTG

LPA_decoy_F			ACAGTAAGCAAAACAAGGTCCA
ON_F2	Fig5E	ON/OT deletion confirmation PCR	TTTTTGCACGTCTGTCTGCC
OT2_R1			AGCCAAGTTGAGTCCTGAGC
OT3_R2			TCAAAAGCAAAGGTCCTGAGAC
OT4_R			TTCTAAGCACGTGGCCATGT
OT6_R1			AGCTCCATGTAGCCTTCTGC
OT12_13_18_26_R2			CGCCATCTGCATCTGTCACA
OT1_F1			AGTTAGCTTGACGCACACCT
OT1_R1			AAGCCTGAGACATTCTGCCC
exon5_probe_MGB_FAM	Fig5F	KIV_2 CNV ddPCR	CTTGGCAGGTTCTTCC
KIV2_copy_F			GACAGAGTTATCGAGGCACATAC
KIV2_copy_R			CTATGCGAGTGTGGTGTCATAG
hRPP30_F			GATTTGGACCTGCGAGCG
hRPP30_R			GCGGCTGTCTCCACAAGT
hRPP30_probe_VIC			CTGACCTGAAGGCTCT

Supplemental References

1. Höijer I, Johansson J, Gudmundsson S, Chin CS, Bunikis I, Häggqvist S, Emmanouilidou A, Wilbe M, den Hoed M, Bondeson ML, et al. Amplification-free long-read sequencing reveals unforeseen CRISPR-Cas9 off-target activity. *Genome Biology*. 2020;21(1):1–19.