A Mini-intronic Plasmid (MIP): A Novel Robust Transgene Expression Vector *In Vivo* and *In Vitro*

Jiamiao Lu[1,2], Feijie Zhang[1,2] and Mark A Kay[1,2]

[1]Department of Pediatrics, Stanford University, Stanford, California, USA; [2]Department of Genetics, Stanford University, Stanford, California, USA

The bacterial backbone (BB) sequences contained within a canonical plasmid DNA dampen exogenous transgene expression by tenfold to 1,000-fold over a period of a few weeks following transfection into quiescent tissues such as the liver. Minicircle DNA vectors devoid of bacterial plasmid backbone sequences overcome transgene silencing providing persistent transgene expression. Because, we recently established that the length rather than sequence of the DNA flanking the transgene expression cassette is the major parameter affecting transgene silencing, we developed an alternative plasmid propagation process in which the essential bacterial elements for plasmid replication and selection are placed within an engineered intron contained within the eukaryotic expression cassette. As with the minicircle vector, the mini-intronic plasmid (MIP) vector system overcomes transgene silencing observed with plasmids but in addition provides between 2 and often 10 times or higher levels of transgene expression compared with minicircle vectors containing the same expression cassette *in vivo* and *in vitro*. These improved plasmids will benefit all studies involving gene transfer/therapy approaches.

Received 10 August 2012; accepted 24 January 2013; advance online publication 5 March 2013. doi:10.1038/mt.2013.33

INTRODUCTION

The term "plasmid" was first introduced by Joshua Lederberg in 1952,[1] and with the advent of recombinant DNA,[2] plasmid DNA-based expression vectors have become an essential tool for both biological discovery and the development of new therapeutics. In terms of their use in gene therapy applications, there are two limitations. The first is the inability to achieve efficient delivery *in vivo* and the second is related to the fact that canonical plasmid-based vectors are incapable of achieving sustained therapeutic levels of transgene expression *in vivo* dramatically limiting their usage. Minicircle DNA vectors devoid of the bacterial backbone (BB) persistently express the transgene at tenfold to 1,000-fold higher levels than that which can be obtained from a conventional plasmid in quiescent tissues *in vivo*.[3–8] In contrast, even though a robust system to produce minicircle DNA vectors has been developed,[9] the production method is still more complicated than conventional plasmid preparation. The mechanism of plasmid-mediated transgene silencing is not fully understood but it is not due to loss of plasmid DNA or changes in DNA methylation.[3,10] Differential chromatin modifications have been noted between plasmid and minicircle DNAs isolated in liver but causal relationships have yet to be established.[11]

We recently recognized that transgene silencing begins to be observed when ~1 kb or more of DNA is placed outside of the transcription unit (between 5′ end of promoter and 3′ end of poly A site), even when the bacterial plasmid DNA sequences are replaced with random DNA sequences.[12] In contrast, if the same extra DNA sequences are moved to within the transcription unit (e.g., made part of the 3′-untranslated region (UTR)), the transgene is not silenced. Therefore, we predicted that whether we could engineer a vector in which the essential plasmid DNA sequences required for bacterial propagation were placed within the transgene expression cassette, we could create a plasmid vector that is not silenced after transfection. Thus, we predicted silencing would be avoided by placing the essential bacterial plasmid DNA sequences required for bacterial propagation into an intron contained within the expression cassette.

The inclusion of an intron may have an added benefit of enhancing transgene expression. It has been reported that many recombinant genes are expressed inefficiently when their introns are removed.[13–15] Splicing of promoter proximal introns has been shown to enhance transcription,[16] and the splicing of terminal introns increases the efficiency of polyadenylation.[17] Introns may increase gene expression through enhancing mRNA export from the nucleus to the cytoplasm.[18] Due to all of these features, we built a new transgene expression system "mini-intronic plasmid (MIP)" that places the bacterial replication origin and selectable marker as an intron in the transgene expression cassette and still keeps the juxtaposition of the 5′ and the 3′ ends of transgene expression cassette as in minicircle. We tested these head to head with canonical plasmid and minicircle DNA vectors *in vitro* and *in vivo*.

RESULTS

Plasmid BB DNA contained in an intron provides sustained transgene expression *in vivo*

This new transgene expression system called "MIP" contains a bacterial replication origin and selectable marker maintaining the juxtaposition of the 5′ and the 3′ ends of transgene expression cassette as in a minicircle. The additional advantages over the minicircle are that the preparation time is shorter (same time as a conventional plasmid), and no antibiotic selection is required.[9]

Correspondence: *Mark A Kay, Departments of Pediatrics and Genetics, Stanford University, 269 Campus Dr. Rm 2105, Stanford, California 94305-5208, USA. E-mail: MARKAY@STANFORD.EDU*

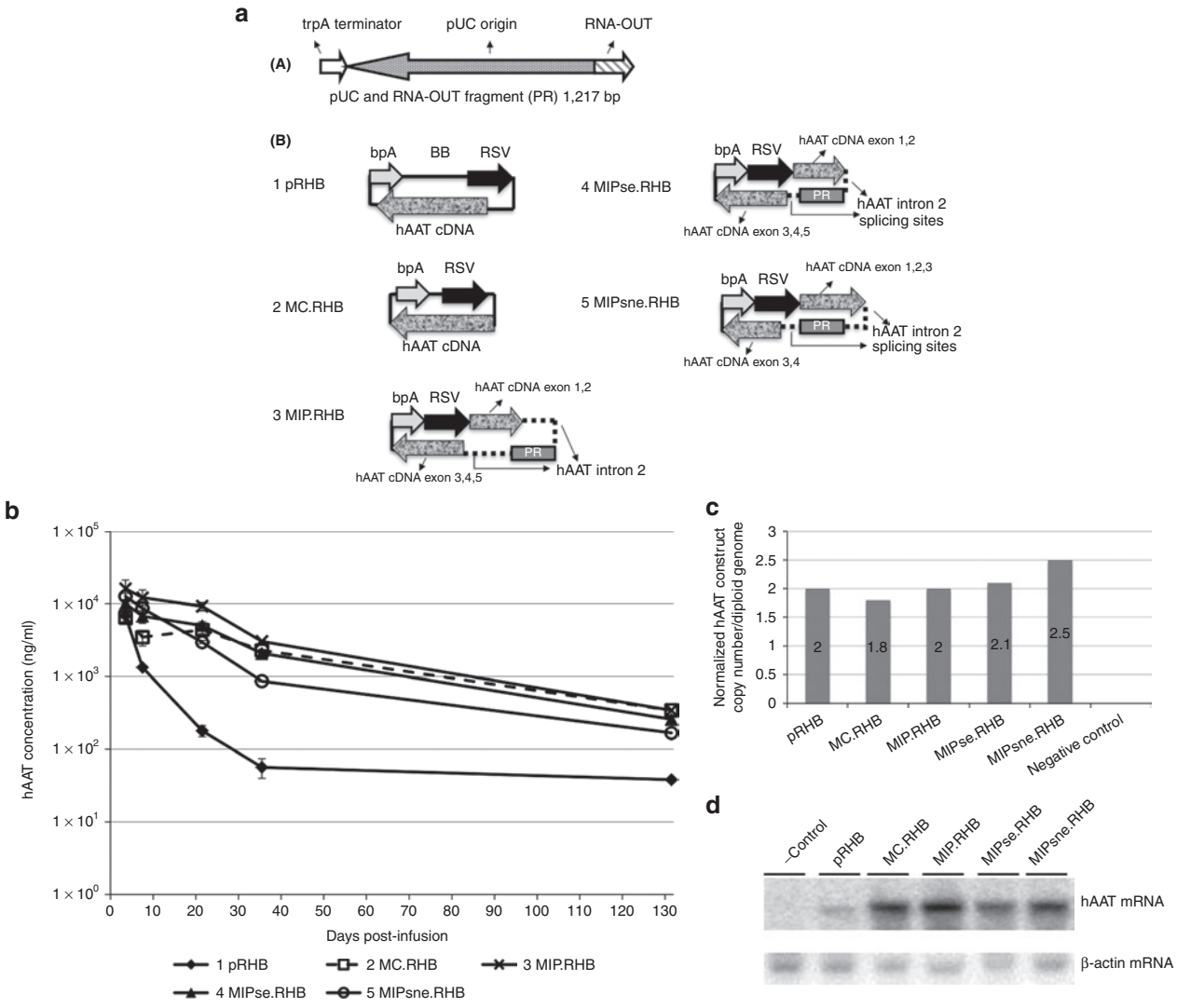
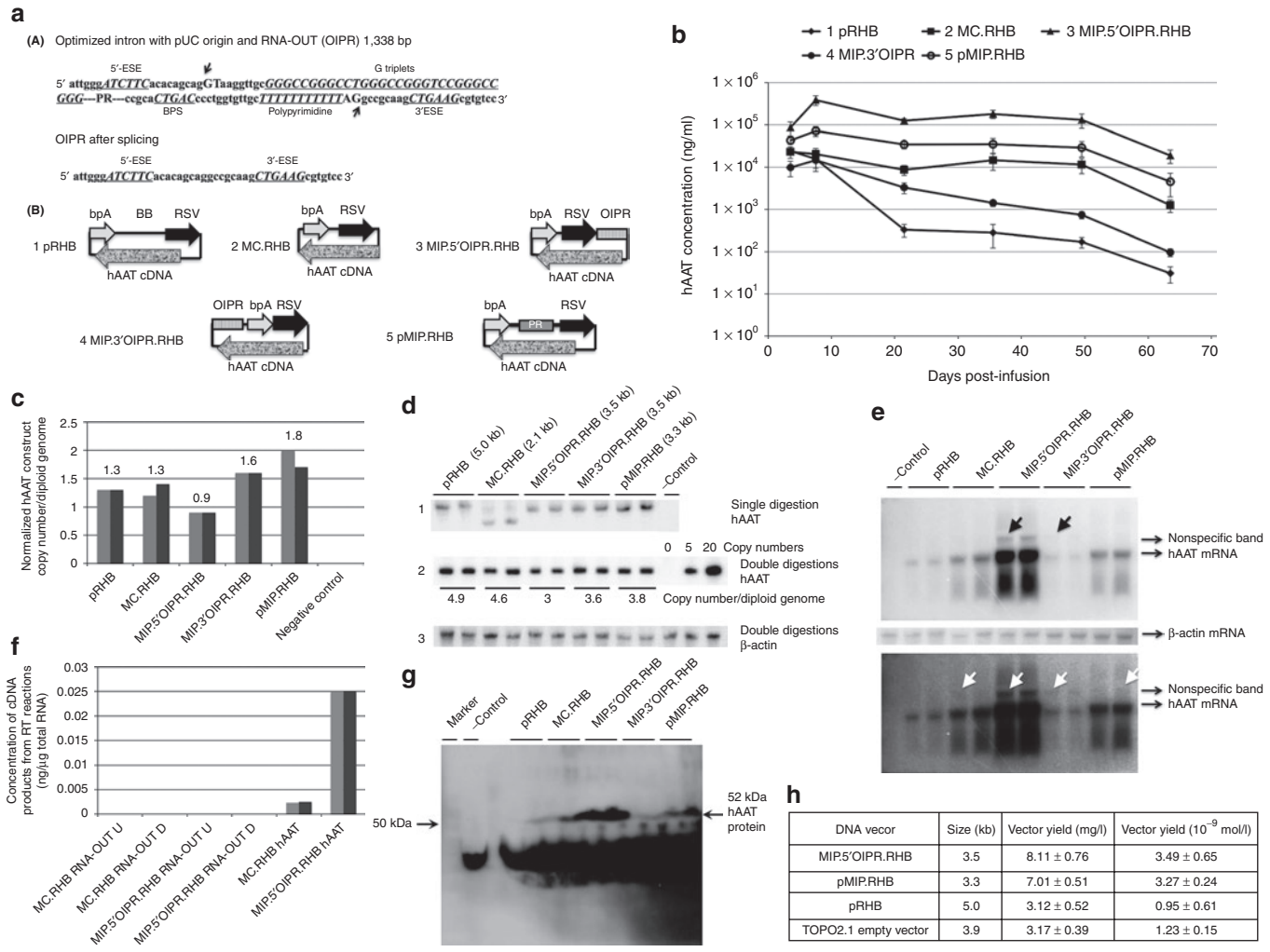


Figure 1 RSV-hAAT expression constructs and transgene expression in mice. (**a**) DNA constructs. (**A**) Schematic structure of the PR sequence. (**B**) Schematic of DNA constructs containing the PR sequence in the intron region. (**b**) Serum hAAT levels at various time points after equimolar infusion of one of the plasmid vectors were infused. (n = 5/group, one animal from each group was terminated on day 35). Error bars represent the SD. (**c**) Copy number of each construct per diploid genome in liver sample 35 days after infusion was determined by quantitative real-time PCR. The mean value was determined by isolating duplicate samples from each animal and was indicated on each bar. For pRHB, 2.0 (2.2, 1.8). For MC.RHB, 1.8 (1.8, 1.8). For MIP.RHB, 2 (2.1, 1.9). For MIP.seRHB, 2.1(2.3, 1.8). For MIPsne.RHB, 2.5 (2.5, 2.5). (**d**) Northern blot data. One 35-day post-infusion liver total RNA sample was extracted from each treated group and probed for hAAT and β-actin transcripts. hAAT, human α1-antitrypsin; RSV, Rous sarcoma virus.

The specific plasmids we created utilized a high copy number pUC origin and antibiotic-free RNA-OUT selectable marker (PR sequence) (**Figure 1a**, **A**; **Supplementary Figure S1a**).[19,20] The 145 nucleotide RNA sequence is transcribed only in prokaryotes and represses expression of a chromosomally integrated constitutively expressed counter-selectable marker (SacB) in the host bacterial strain, allowing plasmid selection when grown in sucrose.[20] In our first set of constructs, the PR sequences were cloned into the native intron 2 of the human α1-antitrypsin (hAAT) minigene (**Figure 1a**, **B**). Additional constructs had the PR sequences inserted into a minimal intron 2 containing the splicing branch point and polypyrimidine tract (**Figure 1a**, **B**). The shortened intron remained positioned at the exons 2 and 3 junction (endogenous location, MIPse.RHB) or placed between exons 3 and 4 (nonendogenous location, MIPsne.RHB) (**Figure 1a**, **B**). These three MIP constructs along with a canonical BB containing plasmid pRHB (silencing control) and BB minus minicircle MC.RHB (non-silencing control) were infused into the livers of 6–8 weeks old C57BL/J female mice by hydrodynamic tail vein injection, respectively, and the level of serum hAAT transgene product was measured at various time points (**Figure 1b**). As we have shown in the past,[3,12] pRHB resulted in high levels of serum hAAT shortly after vector administration, but then declined to very low levels during the next 2–3 weeks. In contrast, the expression from MC.RHB, MIP.RHB, and MIPse.RHB constructs resulted in similar expression profiles (**Figure 1b**). Transgene expression from the MIPsne.RHB vector was slightly reduced compared with MC.RHB and MIP.RHB (**Figure 1b**), but these three vectors expressed at 10 times higher levels than the pRHB canonical plasmid vector.



DNA vecor	Size (kb)	Vector yield (mg/l)	Vector yield (10^{-9} mol/l)
MIP.5′OIPR.RHB	3.5	8.11 ± 0.76	3.49 ± 0.65
pMIP.RHB	3.3	7.01 ± 0.51	3.27 ± 0.24
pRHB	5.0	3.12 ± 0.52	0.95 ± 0.61
TOPO2.1 empty vector	3.9	3.17 ± 0.39	1.23 ± 0.15

Figure 2 RSV-hAAT expression constructs and transgene expression in mice. (**a**) DNA sequences and constructs. (**A**) Structure and sequence of OIPR pre-and post-splicing. 5′-exon splicing enhancer (5′-ESE), G triplets, branch point sequence (BPS), 11 nucleotide polypyrimidine track, and 3′-exon splicing enhancer (3′-ESE) are underlined. Arrows indicate the sequence where splicing takes place. (**B**) Schematics of DNA constructs containing the OIPR sequence. (**b**) Equimolar amounts of DNA plasmids shown in **B** were infused into C57BL/6J mice ($n = 5$/group) and the serum hAAT determined at various time points. Error bars represent the SD. (**c**) Copy number of each construct per diploid genome in infused animal liver samples 63 days post-infusion *via* quantitative real-time PCR. Two treated animals (light and dark gray) from each group were tested and each sample was performed in duplicate. Thus, the mean value was determined by these four readings from each group. The mean value is indicated on each group. For pRHB, 1.3 (1.2, 1.3, 1.3, and 1.3). For MC.RHB, 1.3 (1.3, 1.2, 1.4, and 1.4). For MIP.5′OIPR.RHB, 0.9 (0.9, 0.9, 1.0, and 0.9). For MIP.3′OIPR.RHB, 1.6 (1.6, 1.6, 1.6, and 1.6). For pMIP, 1.8 (1.9, 2.0, 2.0, and 1.4). (**d**) Southern blot analysis of 63 days post-infusion mouse liver genomic DNA samples ($n = 2$/group). Genomic DNA samples were either single-digested with *BamHI* (row 1), or double-digested by *BamHI* and *PmlI* (rows 2 and 3). The size of each vector is marked in bracket. Note that 3.3, 3.5, and 5.0 kb bands are hard to distinguish on this Southern gel. 0, 5, and 20 copies of *BamHI*, *PmlI* double-digested hAAT cDNA were loaded together with 20 μg non-infused negative control genomic DNA as copy number controls. Columns 1 and 2 were probed for hAAT and column 3 was probed for β-actin. (**e**) Northern blot data. Two 63-day post-infusion liver total RNA samples from each group were subjected to northern blot. The hAAT and β-actin transcripts are indicated. A larger, very faint was also detected in animals transfected with MIP.5′OIPR.RHB and MIP.3′OIPR.RHB vectors (black arrows).The bottom panel shows the over-exposed northern blot. As indicated by white arrows, the same larger band was also detected in MC.RHB- and pMIP.RHB-infused groups. (**f**) qPCR quantification of the reverse-transcribed RNA-OUT or hAAT exonic regions ($n = 2$/group). Primer details are in **Supplementary Figure S3b** and Materials and Methods. The RT reactions were performed on MC.RHB- and MIP.5′OIPR.RHB-infused groups by using RNA-OUT–specific and hAAT-specific primers. The cDNA products from RT reactions were further quantified through qPCR. (**g**) Western blot analysis of serum samples. Twenty microliter of 63-day post-infusion animal serum samples from each group ($n = 2$/group) were analyzed. The 52 kDa protein bands corresponding to hAAT proteins are indicated. The 50 kDa protein marker band is indicated with an arrow. (**h**) The yield of MIP plasmid vector and conventional plasmid vectors ($n = 4$/vector). The mol/l yield was calculated as previously described.[9] The linear TOPO2.1 vector (Invitrogen, Grand Island, NY) was digested with *EcoR*I and then self-ligated to form TOPO2.1 empty vector. hAAT, human α1-antitrypsin; MIP, mini-intronic plasmid; qPCR, quantitative PCR; RSV, Rous sarcoma virus; RT, reverse transcription.

In our previous studies,[3,10] we demonstrated that the difference in transgene expression from plasmid and minicircle was not due to differential DNA loss. By quantitative PCR (qPCR), we were able to establish that the MIPs were maintained at the same level as minicircles and plasmids ranging from 1.8 to 2.5 copies per diploid genome, 5 weeks after vector infusion (**Figure 1c**). To establish whether the engineered intron was correctly spliced out of the mRNA, we performed a northern analysis (**Figure 1d**,

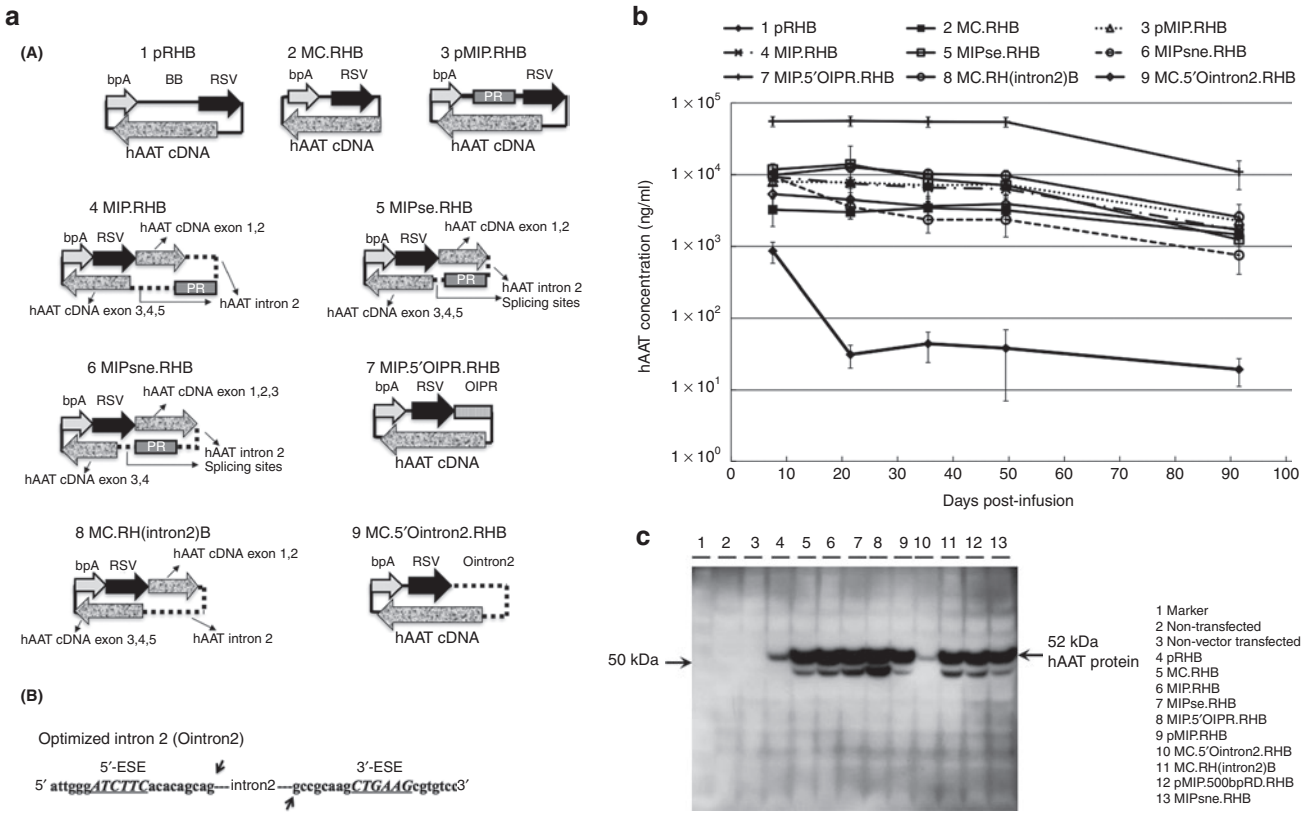


Figure 3 RSV-hAAT expression constructs and transgene expression in mice. (**a**) DNA constructs and sequences. (**A**) Schematic of hAAT expressing DNA constructs. (**B**) Structure and sequence of intron 2 flanked by optimized splicing sites. Arrows indicate the sequence where splicing takes place. (**b**) Equimolar amounts of DNA vectors shown in **Figure 3a**, **A** were infused into mice (n = 5/group). The serum hAAT were determined at various time points. Error bars represent the SD. (**c**) Western blot from HEK293 cell lysates transfected with DNA constructs listed in **Figure 3a**, **A** 24 hours earlier. The 52 kDa hAAT protein bands are indicated. The smaller sized bands likely reflect degraded proteins. The 50 kDa protein marker band is indicated with an arrow. hAAT, human α1-antitrypsin; RSV, Rous sarcoma virus.

Supplementary Figure S2) 5 weeks after liver transfection. We found that all the MIP constructs produced one single-spliced product the same size as the hAAT mRNA produced from pRHB and MC.RHB. This suggests that the presence of an intron containing a PR sequence does not introduce alternative splicing. The strength of the northern blot signal from the use of the different plasmid vectors indicates that the difference in transgene expression was likely due to varied levels of transcription. These data establish that as we predicted, the MIP vectors provide persistent transgene expression *via* mechanisms similar to minicircle vectors.[12]

Optimized intronic sequences contained in a 5′ noncoding exon provides sustained high levels of transgene expression *in vivo*

Now that the proof-of-concept was established, we wanted to produce a MIP vector containing an intronic sequence that could be used with any expression sequence. Thus, we designed vectors that had an optimized intron with a pUC origin and RNA-OUT sequence (OIPR) placed within a noncoding exon. The OIPR sequence contains the pUC origin and RNA-OUT sequence flanked by strong exonic splicing enhancers,[21] G triplets,[22] a tract of 11 continuous polypyrimidines,[23] and a consensus branch point sequence (**Figure 2a**, **A**; **Supplementary Figure S1b**). To do this, we placed the OIPR sequence into a designated 5′ noncoding exon (MIP.5′OIPR.RHB) (**Figure 2a**, **B**) or 3′-UTR (MIP.3′OIPR.RHB) (**Figure 2a**, **B**). pRHB and MC.RHB plasmids were used as silencing and non-silencing controls, respectively. A third control, plasmid pMIP.RHB, comprising the PR sequence (without splicing signals) replacing the conventional BB was constructed to determine whether the presence of the PR sequence had any effect on transgene expression. This PR sequence is 1,217 bp in length, bordering on the length known to cause silencing in previous studies.[12] Of note, it was previously reported that the presence of the RNA-OUT sequence in the plasmid backbone can enhance transgene expression.[20] All of the constructs were infused into mice. The transgene expression data presented in **Figure 2b** showed that the MIP.5′OIPR.RHB expressed the highest levels of transgene product even exceeding minicircle expression (MC.RHB) by ~15 times. Transgene expressed from pMIP.RHB followed the same kinetics as a minicircle vector and was threefold higher than MC.RHB, confirming that the PR sequence itself may improve transgene expression possibly serving a role as an enhancer. Comparing placement of the OIPR sequence as an intron into the 5′ noncoding exon, MIP.5′OIPR.RHB increased transgene expression four- to fivefold higher than when the OIPR sequence was placed outside of the transgene expression cassette (pMIP.RHB), suggesting that RNA splicing of transcripts derived

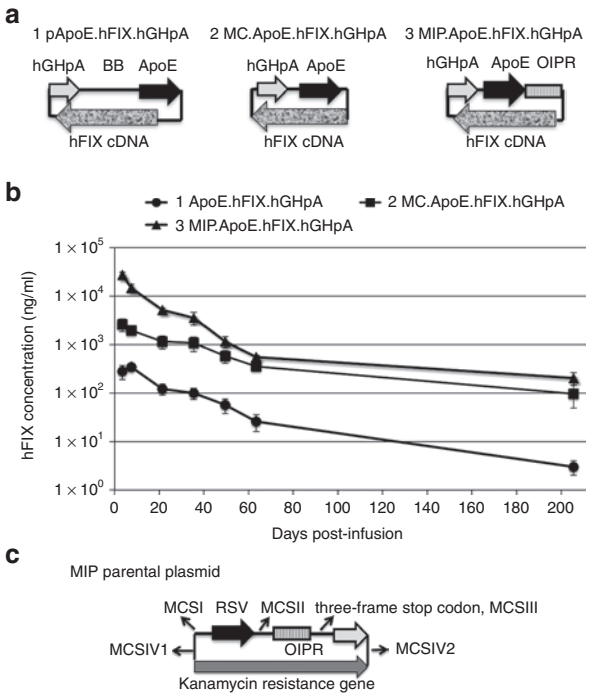


Figure 4 ApoE.hFIX expression constructs and transgene expression in mice. (**a**) Schematic of hFIX expressing DNA constructs. (**b**) Equimolar amounts of DNA constructs shown in **a** were injected into mice (n = 5/group) and the plasma hFIX levels were measured by ELISA at various time points. Error bars represent the SD. (**c**) Structure of the universal MIP cloning vector. ApoE, apolipoprotein E; hFIX, human factor IX; MIP, mini-intronic plasmid.

from MIP.5′OIPR.RHB enhanced expression and/or partial transcriptional silencing of the pMIP.RHB plasmid was contributory. When the OIPR sequence was made as part of the 3′-UTR (MIP.3′OIPR.RHB), the transgene expression was not enhanced probably because changes in the 3′-UTR can affect mRNA stability. Therefore, we abandoned the idea of trying to place the intronic sequences into the 3′-UTR.

To verify that the high level of transgene expression from MIP.5′OIPR.RHB was not due to differential loss or maintenance of episomal DNA vectors, we performed both qPCR assay (**Figure 2c**) and Southern blot analysis (**Figure 2d**) from the livers of treated animals 9 weeks after vector infusion. Single restriction enzyme digestion that cut each vector once verified that the infused vectors remained episomal. The vector copy number per diploid genome varied by twofold or less between all groups as determined in both qPCR (**Figure 2c**) and Southern blot (**Figure 2d**) assays.

To determine the splicing patterns of all tested constructs and to verify whether the different levels of transgene product and mRNAs were correlative, northern blot analyses were performed. As shown in **Figure 2e** and **Supplementary Figure S3a**, the hAAT transcript levels correlated with the levels of protein expression with the MIP.5′OIPR.RHB-infused group having the highest amount of transcript. A trace amount of larger sized band was detected in the MIP.5′OIPR.RHB- and MIP.3′OIPR.RHB-infused groups. However, the same sized band was detected in MC.RHB- and pMIP.RHB-infused samples when the blot was exposed for longer periods (**Figure 2e**, bottom panel), suggesting that this band most likely represented a nonspecific signal and was not the result of an unspliced or mis-spliced mRNA. To prove this point, a series of reverse transcription (RT) and qPCR experiments were performed using various primers derived from exonic hAAT flanking and/or RNA-OUT intronic sequences. The presence of exon–exon but not intronic amplification products shown in **Figure 2f** and **Supplementary Figure S3b** confirmed the absence of an unspliced mRNA. Consistent with this finding was the fact, that no protein isoforms were detected by western blot analyses performed on serum from transfected mice (**Figure 2g**) or transfected cell lysate samples (**Figure 3c**).

The yield of MIP vectors is similar to canonical plasmids

The amount of MIP vectors produced in routine bacterial cultures was compared with routine plasmids. One MIP vector and three bacterial plasmids were propagated and purified from 100 ml cultures four times each (**Figure 2h**). The yield of MIP was at least as high as that obtained with the conventional plasmids.

The presence of PR sequence in the 5′ noncoding exon expressed the highest levels of transgene *in vivo* comparing with conventional plasmid and minicircle vectors

To establish whether the high level of transgene expression from MIP.5′OIPR.RHB vector was due to the presence of the PR sequence itself, an experiment using additional control vectors were tested for transgene expression (**Figure 3a,b**). A MC.RH(intron2)B minicircle vector containing an hAAT intron 2 at the endogenous location without the PR sequence and another minicircle vector, MC.5′Ointron2.RHB containing the hAAT intron 2 in the 5′ noncoding exon flanked by the 5′ exon splicing enhancer and 3′ exon splicing enhancer were constructed. As shown in **Figure 3b**, during this 3-month experiment, the MIP.5′OIPR.RHB-treated mice expressed the transgene at about sevenfold higher levels than the MC.5′Ointron2.RHB-treated mice, and resulted in about 10 times higher levels of transgene expression compared with the corresponding minicircle. Western blot analysis confirmed that the protein produced from these vectors was a single isoform (**Figure 3c**).

To establish the utility of these plasmids, we constructed an MIP vector using a different promoter and transgene. Sustained high levels of human factor IX (hFIX) driven by the ApoE.HCR.hAAT promoter[24] were achieved in transfected mice (**Figure 4a,b**). The MIP.ApoE.hFIX.hGHpA-treated mice had two times higher levels of transgene expression compared with the corresponding minicircle (MC.ApoE.hFIX.hGHpA) treated animals 7 months after infusion.

User friendly improvements in parental MIP design

Antibiotic selection is more commonly used in plasmid DNA propagation so we attempted to replace the RNA-OUT selectable marker in the OIPR sequence with the more commonly used kanamycin resistance gene (**Supplementary Figure S4a**). However, transgene expression was almost two orders of

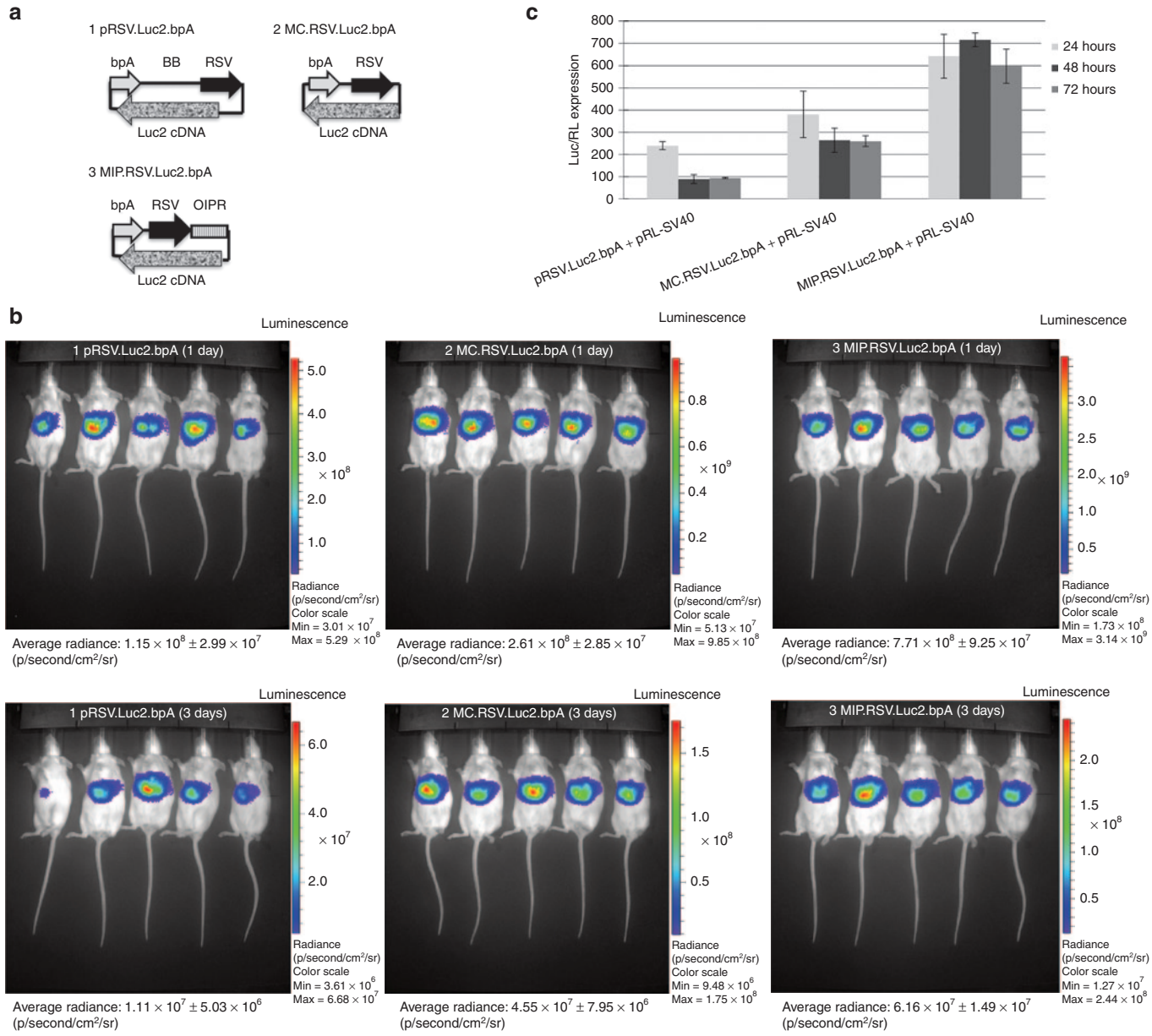


Figure 5 RSV.Luc2 expression constructs and transgene expression in mice and in HEK293 cells. (**a**) Schematic of firefly luciferase2 expressing DNA constructs. (**b**) Equimolar amounts of DNA constructs shown in **a** were injected into 6–8 weeks old FVB mice (n = 5/group) and the *in vivo* liver firefly luciferase2 expression levels were measured at day 1 and day 3 after infusion. The animals were exposed for 1 second. Average radiance (average photons) of each image was marked under the image. Note the luciferase images vary by up to 100 times using a scale that varies from 10^9 to 10^7. (**c**) DNA vectors in **a** were transfected into HEK293 cells with the renilla luciferase (RL) control vector pRL-SV40 at the ratio of 50:1 (n = 3 for each group). Firefly luciferase2 and renilla luciferase signals were analyzed by VERITAS microplate luminometer at 24, 48, and 72 hours after transfection. The firefly luciferase2 signals were normalized to the renilla luciferase signal. RSV, Rous sarcoma virus.

magnitude lower even at the early time points compared with a conventional plasmid (**Supplementary Figure S4b**). This was not surprising because these sequences provide potential alterative splicing sites. Because similar sites exist in other antibiotic selectable markers, we continued to exclusively pursue the RNA-OUT selectable marker in the MIP vector for plasmid propagation.

To make an MIP plasmid suited for routine cloning of any expression cassette, we constructed a universal MIP cloning vector (MIP parental plasmid) containing the OIPR site, Rous sarcoma virus (RSV) promoter, and bpA polyA signal flanked by multiple cloning sites (MCS) (**Figure 4c**). In this vector, MCSI and MCSII can be used to replace the RSV promoter with any other promoter and any transgene-coding sequence can be inserted into MCSIII region. In the small chance of producing unspliced or mis-spliced mRNA, we cloned a three-frame stop codon (TGACTAGCTAA) into MIP parental plasmid between the 3′ end of OIPR sequence and the 5′ end of MCSIII.

While cloning various expression cassettes into the MIP vector, we found that the growth of non-transfected host bacteria on sucrose selection plates increased the number of false-positive clones. To make the cloning step more efficient, we incorporated

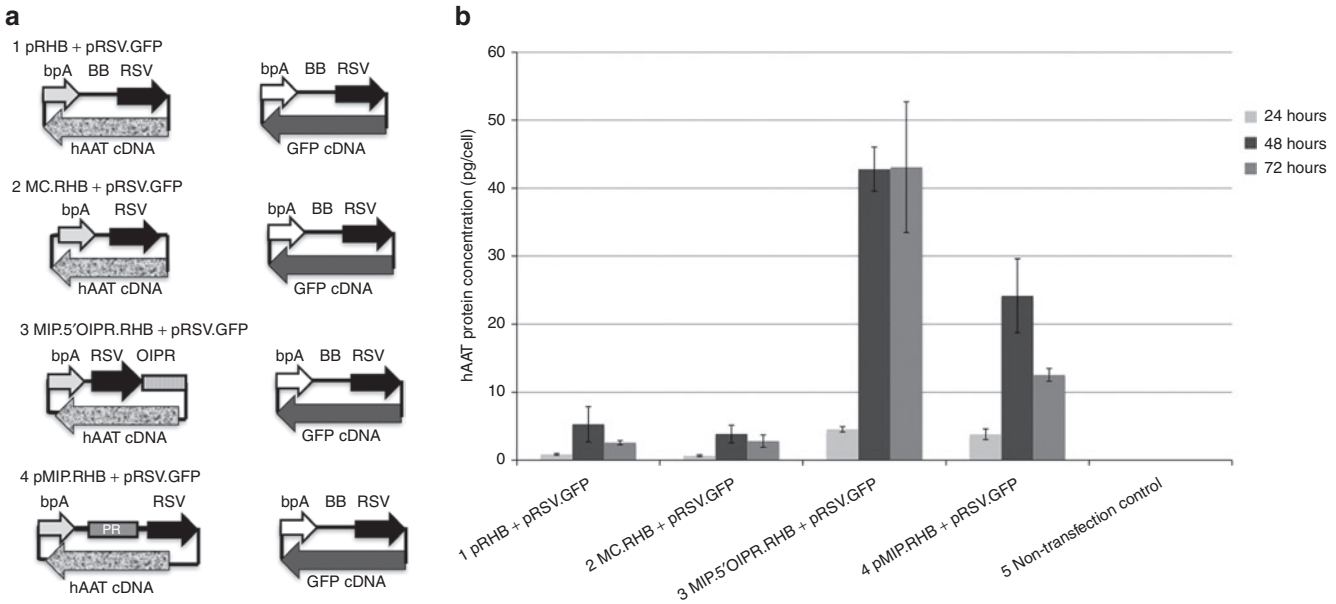


Figure 6 Transgene expression levels in HEK293 cells. (**a**) Schematic of hAAT expressing DNA constructs and GFP construct used for HEK293 transfection. (**b**) hAAT levels in the culture media were quantified at 24, 48, and 72 hours after transfection by ELISA (n = 5/group). The hAAT levels are presented as pg/plated cell. GFP, green fluorescent protein; hAAT, human α1-antitrypsin; RSV, Rous sarcoma virus.

a kanamycin resistance gene in the backbone between MCSIV1 and MCSIV2 restriction sites. This allowed kanamycin selection for the purposes of cloning. The kanamycin gene was then simply removed by a single restriction enzyme digestion of the MCSIV1 and MCSIV2 sites followed by self ligation to produce the final MIP vector. The plasmid vectors derived by this method are stable and easily propagated using RNA-OUT selection.

MIP plasmid vectors produce functional protein products and provide higher levels of transgene expression compared with canonical plasmids *in vivo* and *in vitro*

While the secreted proteins produced by MIP were of the expected size, we wanted to test a third transgene product and unequivocally establish the production of a functional protein. To do this, the firefly luciferase2 gene was cloned into MIP parental plasmid (MIP.RSV.Luc2.bpA) (**Figure 5a**). Vectors, pRSV.Luc2.bpA and MC.RSV.Luc2.bpA were used as plasmid and minicircle controls. Equimolar amounts of these three vectors were transfected into the livers of 6–8 weeks old FVB female mice. Since firefly luciferase may induce an anti-luciferase immune response,[25–27] we measured the firefly luciferase2 expression 1 day and 3 days after infusion. As shown in **Figure 5b**, firefly luciferase2 expressed from MIP vector had a three- and sevenfold higher signal compared with the same expression cassette in a conventional minicircle and conventional plasmid vector, respectively, on day 1 after infusion. The firefly luciferase2 expression levels on day 3 were about 5–10 times lower than on day 1 in all the groups. This rise and fall in luciferase expression was similar to what has been observed when the hepatic luciferase signal expressed from the plasmid vector peaked between 3 and 24 hours after infusion.[25] The decline in the expression was perhaps due to an innate or antigen-dependent immune response.[25–27]

We wanted to compare the relative differences in transgene expression in cultured cells. To do this, equimolar amounts of pRSV.Luc2.bpA, MC.Luc2.bpA and MIP.RSV.Luc2.bpA, and 1/50th of the molar amount of a renilla luciferase expression plasmid, pRL-SV40 (Promega, Madison, WI), were transfected into human HEK293 cells. The firefly luciferase2 expression levels were normalized for renilla luciferase. As shown in **Figure 5c**, MIP vector expressed firefly luciferase2 at about 2.5 and 6 times higher levels than a minicircle and conventional plasmid vector, respectively, 72 hours after transfection. A similar experiment was performed using the hAAT expression vectors. Equimolar amounts of plasmid, minicircle, and MIP vectors expressing hAAT (**Figure 6a,b**) and a green fluorescent protein (GFP) transfection control plasmid (**Supplementary Figure S5**) were transfected into human HEK293 cells (**Figure 6a,b**). The transgene expression levels were analyzed over time. Although GFP expression was similar between groups (**Supplementary Figure S5**), the MIP.5′OIRP.RHB construct expressed hAAT transgene at about 20-fold higher amounts of transgene protein compared with that obtained using a conventional plasmid pRHB or conventional minicircle MC.RHB, 72 hours after transfection, and about threefold higher than pMIP.RHB-transfected cells (**Figure 6a,b**). These two experiments confirm that the same relative MIP-based enhancement in transgene expression was observed both *in vitro* and *in vivo*.

DISCUSSION

The novel MIP vector system is simple to produce and provides superior expression profiles challenging conventional plasmid and minicircle vector design. The insertion of a PR sequence flanked by optimized splicing sequences into the region before the start codon of transgene, allowed us to achieve transgene expression that was even more robust than that achieved by minicircles both *in vivo* and *in vitro*. In this study, we tested three transgenes

by comparing conventional plasmid, minicircle, and new MIP vectors. For all of the three transgenes, MIP vectors had the highest transgene expression levels. As with minicircle, there was some variation in the relative level of transgene expression using different expression cassettes. This may be due to differences in promoter strength, transgene functionality, immune response, and sensitivity of assay used to measure the transgene product. However for all tested transgenes, the general patterns were similar with the MIP vector providing the highest level of transgene expression.

The detailed mechanism of MIP-generated enhanced expression over minicircle DNAs and plasmid DNAs is not clear. Our data support the previously reported finding that the RNA-OUT selectable marker in the vector backbone improved transgene expression.[20] However, since the MIP vectors had higher levels of *in vitro* and *in vivo* transgene expression compared with standard RNA-OUT plasmids, there are likely additional factors influencing the enhanced expression profiles with the MIP vectors. An intron-mediated transcriptional enhancement has been reported for some plant polyubiquitin genes[28] and human ubiquitin C gene.[29] It is possible that the RNA-OUT sequence in the MIP intron functions as a transcriptional enhancer. The inclusion of a spliced promoter proximal intron can enhance transcription levels.[16] Thus, the inclusion of the OIPR into a spliced intronic sequence may have had an additive or even synergistic effect on transgene expression.

Recent studies revealed that plasmid DNA silencing *in vivo* occurs at a nuclear stage that precedes transport of mRNA to the cytoplasm, and dramatic enrichment of H3K27 trimethylation on plasmid sequences was observed compared with minicircle vector.[30] These findings indicate that chromatin-linked transcriptional blockage is a possible mechanism of nuclear silencing.[30] By keeping the bacterial replication origin and selectable marker in the intron, the MIP vector may reduce the amount of H3K27 trimethylation on the vector DNA and prevent chromatin-linked transcriptional blockage.

Other types of DNA-based gene therapy vectors were developed in recent years, such as pFAR[31] and NTC DNA vectors.[20] However, these vectors all have backbone DNA sequences. As previously studied, backbones larger than 1 kb are able to silence transgenes[12] with longer backbones having a greater silencing effect. The backbones in pFAR and NTC DNA vectors make them subject to silencing.

Together, the MIP vectors offer an alternative to the popular plasmid and minicircle vector DNAs that should promote gene transfer studies designed to promote biological discovery and therapeutic applications.

MATERIALS AND METHODS

Materials. The NTC8385-EGFP vector (Nature Technology, Lincoln, NE) was used to isolate PR sequence (containing trpA terminator, pUC bacterial replication origin, and RNA-OUT selectable marker). NTC4862 DH5α competent cells (Nature Technology) were used to produce MIP vectors.

Vector construction. pRHB (pRSV-hAAT-bpA) and MC.RHB were previously described.[3] MC.RH(intron2)B was generated by inserting a RSV-hAAT(exon 1, 2-intron2-exon 3,4,5)-bpA fragment into the minicircle-producing plasmid.[10] The PR sequence was isolated from NTC8385-EGFP vector through PCR reaction with primer pairs: 5′ GCCCGCCTAATGAGCGGGC 3′ and 5′ GGCCGGGGTACCTAATGATTT 3′. A *Sal*I restriction site was incorporated into both the 5′ and 3′ ends of PR sequence and was then ligated to *Xho*I linearized MC.RH(intron2)B to produce vector MIP.RHB. To make MIPse.RHB, a *Ssp*I restriction site was incorporated into the 5′ end of PR sequence and a *Sal*I site was incorporated into the 3′ end of PR sequence. Both *Ssp*I and *Xho*I were used to remove most of the intron2 sequences in MC.RH(intron2)B, and the remaining sequences were then ligated with *Ssp*I- and *Sal*I-digested PR sequence to form MIPse.RHB. The In-Fusion PCR cloning system (Clontech, Mountain View, CA) was used to generate the MIPsne.RHB construct. Primer pairs 5′ TTCCCCTCTCTCCAGGTCTGCCAGCTTACATTTACC 3′ and 5′ GTTGAGCAACCTTACCTTCTGTCTTCATTTTCCAGG3′ (underlined sequences were the homologous sequences used for In-Fusion ligation) were used to PCR amplify the In-Fusion ligation fragment that contains hAAT (exon 4, 5)-bpA-RSV-hAAT (exon 1, 2, 3) from MC.RHB, and primer pairs: 5′ GTAAGGTTGCTCAACCAGCC 3′ and 5′ CTGGAGAGAGGGGAAGGTG 3′ were used to PCR amplify PR sequence with intron2 splicing sites from MIPse.RHB. The two PCR-amplified fragments were then ligated together through the In-Fusion system to generate MIPsne.RHB. The OIPR sequence (**Figure 2a**, **A**) was generated through PCR reactions. An *Nhe*I site was incorporated into the 5′ end of OIPR and a *Spe*I site was incorporated into the 3′ end of OIPR. An *Nhe*I site was incorporated into MC.RHB right before the start codon of the hAATcDNA. Then the *Nhe*I-linearized MC.RHB was ligated with *Nhe*I, *Spe*I double-digested OIPR to build MIP.5′OIPR.RHB. To form MIP.3′OIPR.RHB, a *Pst*I site was incorporated into each end of OIPR. Then a *Pst*I-linearized (after stop codon of hAAT cDNA and before bpA) MC.RHB fragment was used to ligate with *Pst*I-digested OIPR to produce MIP.3′OIPR.RHB. An RSV-hAAT-bpA–linearized fragment was generated by *Spe*I single digestion of MC.RHB. A PR sequence containing *Spe*I on each end was amplified through PCR reaction. These two *Spe*I-digested fragments were ligated together to make the pMIP.RHB construct. The MC.5′Ointron2.RHB vector was built through In-Fusion ligation. Primer pairs: 5′ AGCAGGTAAGGTTGCTCAACCAGCC 3′ and 5′ TTCAGCTTGCGGCCTGGAGAGAGGGGAAGGTG 3′ were used to amplify hAAT intron 2, and primer pairs: 5′ AGGCCGCAAGCTGAAGCGT 3′ and 5′ GCAACCTTACCTGCTGTGTG 3′ were used to generate hAAT-bpA-RSV sequence that containing optimized splicing sites before the hAAT start codon. The In-Fusion ligation of these two fragments resulted in the MC.5′Ointron2.RHB vector.

MIP vector production. Ligation reactions were transformed into NTC4862 DH5α competent cells. Cells were then plated on 6% sucrose solid media (Nature Technology's NTC vector user manual) and propagated at 30 °C for 24–48 hours. Colonies were picked up for mini-prep; 100 µl bacterial culture of positive mini-prep was then inoculated into 100 ml of 6% sucrose liquid media and incubated for 16–18 hours at 37 °C. The overnight culture carries MIP vector was then produced through conventional plasmid vector preparing procedure by using QIAGEN plasmid MAXI prep kit (QIAGEN, Valencia, CA).

Animal studies. All animal procedures were performed under the guidelines set forth by the National Institutes of Health and after approval from the Stanford Animal Care Committee. Six to eight weeks old female C57BL/6J mice purchased from Jackson Laboratory (Bar Harbor, ME) were used for DNA vector infusion. To ensure that equimolar amount of DNA was injected into each animal, 3.63 µg/kb DNA was used for various sized constructs in **Figures 1b,2b** and **4b**. A total of 1.8 µg/kb DNA was used for various sized constructs in **Figures 3b** and **5b**. Each DNA construct was diluted into 1.8 ml of 0.9% NaCl for each animal, and was delivered through hydrodynamic tail vein injection. Five animals were tested for each DNA construct in each tested experimental group. After DNA infusion, blood samples were collected periodically by a retro-orbital technique. The serum hAAT and the plasma hFIX levels were quantified by ELISA as described earlier.[32]

Southern blot analysis. The salt-out procedure was used to extract liver genomic DNA. Two representative animals from each treated group were selected. The liver genomic DNA of each animal was then digested with either *BamH*I alone or *BamH*I and *Pml*I overnight at 37 °C. *BamH*I cut the expression cassette once. *BamH*I, *Pml*I double digestion cut the expression cassette twice. The digested genomic DNA was then extracted using phenol-chloroform. Twenty microgram of phenol-chloroform extracted genomic DNA of each animal was loaded on 1% agar gel. Various amounts of the AAT cDNA were also digested with *BamH*I and *Pml*I, and then mixed with 20 µg non-infused mouse liver genomic DNA in each lane as copy number control. Southern blot membrane was hybridized with [P-32] dCTP-labeled *BamH*I and *Pml*I double-digested 600 bp hAAT cDNA fragment, and [P-32] dCTP-labeled 300 bp β-actin (PCR product by using forward primer 5′ ACGCGTCCAATTGCCTTTCT 3′ and reverse primer 5′ CTCGA GGTTGAAGGTCTCAA 3′).The copy number calculation method was as previously described.[12] The Southern blot signal strength was measured using Quantity One software (Bio-Rad Laboratories, Hercules, CA).

qPCR analysis. 100 ng of *BamH*I-digested genomic DNA (**Figure 1c**) and *BamH*I, *Pml*I double-digested genomic DNA (**Figure 2c**) from each sample was used as the template for qPCR. One (for **Figure 1c**) or two (for **Figure 2c**) animal samples from each injection group were selected and two 20 µl reactions were performed for each animal sample. Various copy numbers (2×10^8 copies to 20 copies) of double-digested standard vector DNA along with 100 ng of non-infused control genomic DNA per reaction were used to calculate a copy number standard curve. Forward primer: 5′ AAGGCAAATGGGAGAGACCT 3′ and reverse primer: 5′ TACCCAGCTGGACAGCTTCT 3′ oligos were used to amplify 150 bp fragment from hAAT cDNA region. Forward primer: 5′ TTGCTGACAGGATGCAGAAG 3′ and reverse primer: 5′ TGATCCAC ATCTGCTGGAAG 3′ oligos were used to amplify 150 bp fragment from β-actin as loading control. The tested transgene signal was then normalized to the β-actin signal. The mass of a single diploid copy of mouse genome is 5.88 pg. Thus, 100 ng of genomic DNA contains 17,007 copies of diploid genome (1×10^6/5.88 pg). The average transgene copy number in 100 ng of genomic DNA from each group was then divided by 17,007 to achieve the transgene construct copy number in each cell. All calculations were based on methods described by AppliedBiosystems-method, http://www6.appliedbiosystems.com/support/tutorials/pdf/quant_pcr.pdf and by using internet tool, http://www.uri.edu/research/gsc/resources/cndna.html. qPCR was performed using Corbett Research RG6000 PCR instrument (Corbett Research, Mortlake, Australia).

Northern blot analysis. hAAT transcript was detected by standard northern blotting methods, using total liver RNA isolated by TRIzol (Invitrogen, Carlsbad, CA). A total of 20 µg liver RNA was used for each sample. Two identical northern blots were prepared for each experiment. One northern blot was probed with a 220 bp long [P-32] dCTP-labeled hAAT cDNA probe to detect hAAT mRNA; 220 bp hAAT cDNA probe was generated through PCR by using primers: 5′ GCCGTCTTCTGTCTCGTGG 3′ and 5′ AAGAAGATATTGGTGCTGTTGG 3′. The other northern blot was probed with [P-32] dCTP-labeled β-actin probe to detect β-actin as loading control.

RT reaction and following qPCR reaction. RT reaction was performed according to the manual of SuperScript III First-Strand Synthesis System for RT-PCR from Invitrogen. cDNA synthesis was performed in the first step using 5 µg DNaseI-treated liver total RNA sample primed with gene-specific primer (RNA-OUT U, RNA-OUT D, or hAAT D). The sequence of RNA-OUT U primer is 5′ GTAGAATTGGTAAAGAGAGTCG 3′. The sequence of RNA-OUT D primer is 5′ GGTACCTAATGATTTTTATCAAAA 3′. The sequence of hAAT D primer is 5′ TTATTTTTGGGTGGGATTCACC 3′. The synthesized RNA-OUT cDNA was further quantified through qPCR with primers RNA-OUT qPCR U (5′ GGTAAAGAGAGTCGTGTAAAAT 3′) and RNA-OUT qPCR D (5′ GATTTTTATCAAAATCATTAAGTTAA 3′). The synthesized hAAT cDNA was further quantified through qPCR with primers hAAT qPCR U (5′ AAGGCAAATGGGAGAGACCT 3′) and hAAT qPCR D (5′ TACCCAGCTGGACAGCTTCT 3′). Linearized RNA-OUT DNA templates at the concentrations of 0.1, 0.01, 0.001, 0.0001, and 0.00001 ng/µl were used to set up qPCR standard. The results were further normalized by the amount of total RNA used for RT reaction.

Western blot analysis. hAAT expression in mouse serum and the HEK293 cell line was validated by western blotting using a monoclonal anti-hAAT antibody (Abcam, Cambridge, MA). Serum samples (20 µl/sample) from treated animals 9 weeks after vector infusion along with non-treated animal sera were used to perform western blot. 8×10^5 HEK293 cells were seeded in each well of a 6-well tissue culture plate 1 day before transfection. A total of 4 µg of each DNA vector was transfected into each well according to the Lipofectamine 2000 Transfection Reagent manual from Invitrogen. Cell lysates were then harvested 24 hours after transfection and used for the study. hAAT protein was visualized using the monoclonal antibody (Abcam) at a 1:4,000 dilution.

Firefly luciferase in vivo imaging. D-luciferin (firefly) potassium salt was dissolved in Dulbecco's phosphate-buffered saline to make 15 mg/ml injection solution. The luciferin injection solution was infused into tested mice through intraperitoneal injection. A dose of 10 µl/g mouse weight is used. The luciferase signal was examined at 10 minutes after injection by Xenogen IVIS 200 spectrum (Xenogen, Alameda, CA). The luciferase2 images in **Figure 5b** were taken under the following settings: emission filter = open, excitation filter = block, Bin: (M)8, field of view: 22.4, f4, 1 second. Quantitation of signal was calculated using the Living Image 4.3 Software (Caliper Life Sciences, Mountain View, CA).

HEK293 cell transfection. 1.6×10^5 HEK293 cells were seeded in each well of a 24-well tissue culture plate 1 day before transfection. A total of 500 µl media was used in each well. For the experiments shown in **Figure 6b**, equimolar amount of each DNA vector was cotransfected with pRSV.GFP DNA vector at the ratio of 3:1 into each well according to the Lipofectamine 2000 Transfection Reagent manual from Invitrogen. GFP expression was checked with a ultraviolet microscope 24, 48, and 72 hours after transfection. Cell culture media was changed every 24 hours after transfection and the hAAT transgene levels in the media were tested through ELISA. The resulting total hAAT levels every 24 hours were divided by total cell number plated into each well. For the experiments shown in **Figure 5c**, the same molar amount of each DNA vector was cotransfected with pRL-SV40 DNA vector at the ratio of 50:1 into each well. Transfected HEK293 cells were treated by Dual-Luciferase Reporter 1000 Assay System (Promega). The firefly luciferase2 and renilla luciferase signals were detected by VERITAS microplate luminometer (Turner BioSystems, Sunnyvale, CA) 24, 48, and 72 hours after transfection.

SUPPLEMENTARY MATERIAL

Figure S1. Detailed sequence for PR and OIPR fragments. A) PR sequence. The PR sequence was isolated from NTC8385-EGFP vector (Nature Technology Corporation). Capitalized sequence "CC" (32–33) was mutated from original sequence "gg" to prevent formation of a potential alternative splice site. The trpA prokaryotic terminator (underlined sequence 1–30) in this PR sequence prevents prokaryotic transcription during selection. The RNA-OUT selectable marker (capitalized sequence 1066–1210) expresses a 145 bp RNA-OUT antisense RNA, which represses the expression of a counter-selectable marker, SacB from the host bacteria. SacB gene encodes levansucrase, which is toxic to bacteria in the presence of sucrose. Thus vectors with the RNA-OUT selectable marker enable the host bacteria to grow on media in the presence of sucrose [20]. B) OIPR sequence. 5′ESE is indicated as capitalized sequence 7–12. G triplets are capitalized sequence 32–65. BPS is capitalized sequence 1288–1292. 11 nucleotides polypyrimidine sequence is capitalized from 1305–1315. 3′ESE is indicated as capitalized sequence 1326–1331.

Figure S2. Complete northern blot images for **Figure 1d**.

Figure S3. Complete northern blot images and description of the RNA-OUT primers used for RT and qPCR reactions.

Figure S4. RSV.hAAT expression constructs and transgene expression in mice.

Figure S5. GFP expression as transfection efficiency control.

REFERENCES

1. Lederberg, J (1952). Cell genetics and hereditary symbiosis. *Physiol Rev* **32**: 403–430.
2. Kopecko, DJ and Cohen, SN (1975). Site specific recA–independent recombination between bacterial plasmids: involvement of palindromes at the recombinational loci. *Proc Natl Acad Sci USA* **72**: 1373–1377.
3. Chen, ZY, He, CY, Ehrhardt, A and Kay, MA (2003). Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression *in vivo*. *Mol Ther* **8**: 495–500.
4. Chen, ZY, He, CY and Kay, MA (2005). Improved production and purification of minicircle DNA vector free of plasmid bacterial sequences and capable of persistent transgene expression *in vivo*. *Hum Gene Ther* **16**: 126–131.
5. Bigger, BW, Tolmachov, O, Collombet, JM, Fragkos, M, Palaszewski, I and Coutelle, C (2001). An araC-controlled bacterial cre expression system to produce DNA minicircle vectors for nuclear and mitochondrial gene therapy. *J Biol Chem* **276**: 23018–23027.
6. Darquet, AM, Cameron, B, Wils, P, Scherman, D and Crouzet, J (1997). A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. *Gene Ther* **4**: 1341–1349.
7. Mayrhofer, P, Blaesen, M, Schleef, M and Jechlinger, W (2008). Minicircle-DNA production by site specific recombination and protein-DNA interaction chromatography. *J Gene Med* **10**: 1253–1269.
8. Schakowski, F, Gorschlüter, M, Buttgereit, P, Märten, A, Lilienfeld-Toal, MV, Junghans, C *et al.* (2007). Minimal size MIDGE vectors improve transgene expression *in vivo*. *In Vivo* **21**: 17–23.
9. Kay, MA, He, CY and Chen, ZY (2010). A robust system for production of minicircle DNA vectors. *Nat Biotechnol* **28**: 1287–1289.
10. Nakai, H, Yant, SR, Storm, TA, Fuess, S, Meuse, L and Kay, MA (2001). Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction *in vivo*. *J Virol* **75**: 6969–6976.
11. Riu, E, Chen, ZY, Xu, H, He, CY and Kay, MA (2007). Histone modifications are associated with the persistence or silencing of vector-mediated transgene expression *in vivo*. *Mol Ther* **15**: 1348–1355.
12. Lu, J, Zhang, F, Xu, S, Fire, AZ and Kay, MA (2012). The extragenic spacer length between the 5′ and 3′ ends of the transgene expression cassette affects transgene silencing from plasmid-based vectors. *Mol Ther* **20**: 2111–2119.
13. Brinster, RL, Allen, JM, Behringer, RR, Gelinas, RE and Palmiter, RD (1988). Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci USA* **85**: 836–840.
14. Palmiter, RD, Sandgren, EP, Avarbock, MR, Allen, DD and Brinster, RL (1991). Heterologous introns can enhance expression of transgenes in mice. *Proc Natl Acad Sci USA* **88**: 478–482.
15. Lacy-Hulbert, A, Thomas, R, Li, XP, Lilley, CE, Coffin, RS and Roes, J (2001). Interruption of coding sequences by heterologous introns can enhance the functional expression of recombinant genes. *Gene Ther* **8**: 649–653.
16. Furger, A, O'Sullivan, JM, Binnie, A, Lee, BA and Proudfoot, NJ (2002). Promoter proximal splice sites enhance transcription. *Genes Dev* **16**: 2792–2799.
17. Niwa, M, Rose, SD and Berget, SM (1990). *In vitro* polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev* **4**: 1552–1559.
18. Reed, R and Hurt, E (2002). A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* **108**: 523–531.
19. Carnes, AE, Luke, JM, Vincent, JM, Anderson, S, Schukar, A, Hodgson, CP *et al.* (2010). Critical design criteria for minimal antibiotic-free plasmid vectors necessary to combine robust RNA Pol II and Pol III-mediated eukaryotic expression with high bacterial production yields. *J Gene Med* **12**: 818–831.
20. Luke, J, Carnes, AE, Hodgson, CP and Williams, JA (2009). Improved antibiotic-free DNA vaccine vectors utilizing a novel RNA based plasmid selection system. *Vaccine* **27**: 6454–6459.
21. Fairbrother, WG, Yeh, RF, Sharp, PA and Burge, CB (2002). Predictive identification of exonic splicing enhancers in human genes. *Science* **297**: 1007–1013.
22. McCullough, AJ and Berget, SM (1997). G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection. *Mol Cell Biol* **17**: 4562–4571.
23. Coolidge, CJ, Seely, RJ and Patton, JG (1997). Functional analysis of the polypyrimidine tract in pre-mRNA splicing. *Nucleic Acids Res* **25**: 888–896.
24. Miao, CH, Ohashi, K, Patijn, GA, Meuse, L, Ye, X, Thompson, AR *et al.* (2000). Inclusion of the hepatic locus control region, an intron, and untranslated region increases and stabilizes hepatic factor IX gene expression *in vivo* but not *in vitro*. *Mol Ther* **1**: 522–532.
25. Wilber, A, Frandsen, JL, Wangensteen, KJ, Ekker, SC, Wang, X and McIvor, RS (2005). Dynamic gene expression after systemic delivery of plasmid DNA as determined by *in vivo* bioluminescence imaging. *Hum Gene Ther* **16**: 1325–1332.
26. Limberis, MP, Bell, CL and Wilson, JM (2009). Identification of the murine firefly luciferase-specific CD8 T-cell epitopes. *Gene Ther* **16**: 441–447.
27. Jeon, YH, Choi, Y, Kang, JH, Kim, CW, Jeong, JM, Lee, DS *et al.* (2007). Immune response to firefly luciferase as a naked DNA. *Cancer Biol Ther* **6**: 781–786.
28. Sivamani, E and Qu, R (2006). Expression enhancement of a rice polyubiquitin gene promoter. *Plant Mol Biol* **60**: 225–239.
29. Bianchi, M, Crinelli, R, Giacomini, E, Carloni, E and Magnani, M (2009). A potent enhancer element in the 5′-UTR intron is crucial for transcriptional regulation of the human ubiquitin C gene. *Gene* **448**: 88–101.
30. Maniar, LEG, Maniar, JM, Chen, Z-Y, Lu, J, Fire, AZ and Kay, MA (2012). Minicircle DNA vector achieve sustained expression reflected by active chromatin and transcriptional level. *Mol Ther* **21**: 131–138.
31. Marie, C, Vandermeulen, G, Quiviger, M, Richard, M, Préat, V and Scherman, D (2010). pFARs, plasmids free of antibiotic resistance markers, display high-level transgene expression in muscle, skin and tumour cells. *J Gene Med* **12**: 323–332.
32. Yant, SR, Meuse, L, Chiu, W, Ivics, Z, Izsvak, Z and Kay, MA (2000). Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat Genet* **25**: 35–41.