SUPPLEMENTARY INFORMATION

for

TALEN/CRISPR-mediated engineering of a promoterless anti-viral RNAi hairpin into an endogenous miRNA locus

Elena Senís, Stefan Mockenhaupt, Daniel Rupp, Tobias Bauer, Nagarajan Paramasivam, Bettina Knapp, Jan Gronych, Stefanie Grosse, Marc P. Windisch, Florian Schmidt, Fabian J. Theis, Roland Eils, Peter Lichter, Matthias Schlesner, Ralf Bartenschlager, and Dirk Grimm

CONTENTS:

SUPPLEMENTARY DISCUSSION

During analysis of our shmiRHCV318-engineered Huh7 clones, we made three surprising observations. Below, we will discuss what we consider as the most likely hypotheses that can explain these findings.

A. On-target effects in shmiRHCV318-engineered clones

We noted that cellular clones harboring the integrated shmiRHCV318 hairpin displayed some degree of heterogeneity, most of which can be readily explained. In fact, out of the seven clones that we analyzed in more detail, three had no mutations at all (T2 4.37, T2 31.3, U6 6.30), neither in the targeted miR-122 locus nor in the integrated shmiRNA hairpin, which is the ideal outcome. Also important to point out is that the G-to-C conversion in the edited allele of clone U6 7.16 results from the use of a PAM-mutated repair template and was thus expected (as also noted in the legends to Supplementary Figure S7).

Furthermore, three other mutations shown in Figure 4 - (i) a 1 bp insertion in clone TS 30.20, (ii) a 5 bp deletion in clone U6 7.16, and (iii) a 4 bp deletion in clone U6 6.21 - all occurred in the miR-122 allele of these heterozygous clones that was devoid of the shmiRHCV318 hairpin. It is thus very likely that these are all normal indels that resulted from faulty non-homologous end-joining (NHEJ) and that are prone to happen in any TALEN/CRISPR reaction, regardless of the presence or absence of a repair template. Accordingly, we believe that these alleles were cut and then NHEJ-mutated, while the other allele has properly integrated the foreign DNA. Additional strong evidence is that all three mutations were found in the exact position that was expected for the specific TALEN or gRNA, implying they are indeed NHEJ by-products. Hence, this particular type of heterogeneity, *i.e.*, on-target indel mutations in alleles that were cleaved but failed to take up a repair template, is normal and must be anticipated.

This leaves only three mutations in our seven clones that require additional explanation. We speculate that two of them - (i) the 66 bp insertion in both alleles of clone U6 20.16, and (ii) the 13 bp deletion in clone TS 30.20 - can be explained by the fact that in both cases, the cut site of the respective nuclease (Cas9 for clone U6 20.16, and TALEN Seed for clone TS 30.20) was fully conserved in the repair template that we used for homology-directed miR-122 editing. We thus assume that the nucleases may have cleaved not only the genomic locus, but also the repair template, causing inadvertent NHEJ-mediated mutations that were subsequently incorporated into the edited miR-122 allele. Indeed, such a mechanism - whereby the repair DNA was already mutated prior to homologous recombination - would readily explain the curious observation that both alleles in clone U6 20.16 carried the exact same 66 bp insertion, and it could likewise account for the 13 bp deletion in one allele of clone TS 30.20. We note that the origin of the 66 bp fragment remains unclear since it did not align with any of the DNA sequences used in the process, including the donor template and the miR-122 target region. Alternatively, it is also possible that two subsequent cleavage events happened, of which the first triggered integration of the shmiRHCV318 hairpin, while the second resulted in re-cleavage of the target site (which in the case of U6 20.16 and TS 30.20 was conserved in the repair DNA, see above). Consequently, such a two-step

cleavage mechanism can explain both, the correct integration of the foreign DNA (step #1) and its subsequent mutation due to re-cutting and NHEJ (step #2).

The remaining third mutation - a 195 bp deletion in the edited allele of clone U6 6.21 that had resulted in loss of the entire shmiRHCV318 hairpin, but not the floxed Neo^R cassette - is perhaps most surprising since it occurred 50 nt upstream of the predicted gRNA cleavage site. Therefore, an explanation could be the secondary structure of the shmiRHCV318 hairpin, which caused intra-molecular recombination and subsequent loss of the ectopic stem-loop structure prior to or during chromosomal integration.

In conclusion, all mutations in the miR-122 locus that we observed probably fall into three categories: (i) normal indels that occurred in the non-edited allele due to NHEJ, (ii) artefacts from unintentional mutation of the repair DNA prior to or after integration, or (iii) deletion of the foreign RNAi hairpin due to secondary structure-mediated recombination. We reckon that the first two side effects may be improved via strategies already mentioned in the main text, *e.g.*, by applying conditions that shift the balance from NHEJ to homology-directed repair, by using repair templates that are resistant to nuclease cleavage, and/or by designing homology integration strategies such that correct gene editing eliminates nuclease binding or cleavage sites. To also reduce the risk of recombination and deletion of the exogenous hairpin, it should help to further minimize its secondary structure where possible without affecting functionality; besides, this side effect seems relatively rare based on our data at this point.

B. Perturbation of mature miR-122 steady-state levels

Even in our lead clone T2 31.3, the levels of mature miR-122 are roughly two-fold lower than in wildtype Huh7 cells, as consistently detected by Northern blotting (Figure 5B) and qRT-PCR analyses (Figure 5D). Importantly, this slight reduction of miR-122 levels may not bear functional consequences for the cells. This is implied by our findings that a miR-122-sensitive luciferase reporter was knocked down to a comparable extent in both, clone T2 31.3 and its parental Huh7 cells (Figure 5E). Moreover, gene and miRNA expression profiles in clone T2 31.3 were largely similar to the wildtype cells (Figure 6), solidifying that the two-fold reduction of mature miR-122 in our engineered cells was well tolerated. We thus conclude that the slight perturbance of miR-122 levels in our lead cellular clone was unproblematic.

Still, it remains interesting to speculate which mechanisms may account for the miR-122 perturbation. As a matter of fact, it is still largely unclear and widely debated at which stage(s) in the multi-step RNAi pathway steady-state miRNA levels are primarily controlled. One exciting hypothesis is that nuclear processing of the pri(mary)-miRNA by the Drosha/DGCR8 microprocessor is the key regulatory step and as such even more important than transcription of the primary miRNA. This has led, for instance, Conrad *et al*. to recently postulate that the microprocessor is "the main hub for diversified miRNA output" which uncouples cellular miRNA from gene expression (1). This is in line with numerous prior studies which reported evidence for co-transcriptional Drosha processing of pri-miRNA transcripts into pre-miRNAs, which are then exported into the cytoplasm for further processing by Dicer (*e.g.*, (2,3)). A possible consequence of such co-transcriptional Drosha cleavage is the generation of a free 5' end which is a

substrate for the Xrn2 exonuclease, resulting in destabilization of RNA polymerase II and transcription termination, and/or nucleolytic degradation of the 3'-cutoff products from Drosha cleavage. This, in turn, can readily help to explain our findings that the most 3' hairpin - always miR-122 in our strategy - is slightly reduced in the presence of additional hairpins located 5'. In contrast, the most 5' hairpin - shmiRhAAT - is least affected by multiplexing (Figure 2C). Congruent with this, many groups have reported substantial differences of up to two orders of magnitude between the steady-state levels of multiple RNAi hairpins that are co-encoded in the same primary transcript, as part of either a natural miRNA or a synthetic shRNA cluster (*e.g.*, (4-6)).

Importantly, the order of RNAi hairpins in a cluster is, by far, not the only parameter that determines miRNA/shRNA steady-state levels. Instead, equally essential are specific sequence elements within the RNAi hairpin, its secondary structure, its position with the transcript, or the genomic context, to name a few. In addition, Haar *et al*. recently showed that one miRNA hairpin can directly influence the processing of another within a cluster, in a sequence-independent but position-dependent manner (7)). As a final example illustrating the complexity of miRNA expression and processing, several groups documented that the polycistronic miR-17~92 cluster (comprising six miRNAs) assumes a compact, higher-order tertiary RNA structure, whereby those miRNAs that are internalized within the core are processed less efficiently than those on the surface. Thus, Chaulk, Yang and Chakraborty plus respective colleagues postulated that the microprocessor can better access these exposed cluster members, whose processing is followed by Xrn2 cleavage (4-6). This, in turn, provides access to the inner core miRNAs, hence explaining the observed differences in expression levels.

Notably, our data show a very high correlation between the levels of pre-miR-122 and mature miR-122 in each cell line, with the exception of clone TS 30.20 where the pre-miR-122 primers did not bind due to the mutations. We thus conclude that cytoplasmic Dicer processing is not perturbed in our engineered cells, rather suggesting an effect at the level of pri-miR-122 expression and/or nuclear processing along the mechanisms outlined above. To further dissect the underlying processes, it would, for instance, be helpful to experimentally determine the secondary structure of an engineered miR-122 locus using SHAPE technology (8,9), to see whether it matches our predictions in Supplementary Figure S2. Combined with variations in the structure of the miR-122 scaffold and of its position within the *hcr* locus (5' or 3' of miR-122, and at different distances), this should eventually help to fine-tune and optimize the expression/processing of both, the ectopic and the endogenous hairpins.

C. Dysregulation of miRNA and gene profiles

A third curious observation was that clone U6 20.16 differed significantly from the other two engineered cell lines, by showing distinct expression of various miRNAs (next to miR-122) and genes. A first, straight-forward explanation could be wide-spread occurrence of off-target mutations that have disrupted the cognate coding sequences. While we cannot formally exclude this possibility owing of the lack of whole genome sequencing data for this particular clone, we consider this mechanism rather unlikely, at least as

a sole explanation. This is because we would expect the majority of these mutations to ablate expression of the affected gene or miRNA, which contradicts our data. Instead, several miRNAs are actually upregulated in clone U6 20.16, including various members of the let7 family, or miR-20a-5p and miR-17-5p. Likewise, we noted large numbers of genes whose expression is in- rather than decreased. Besides, our prior work yielded no experimental evidence for off-targeting by the same anti-miR-122 gRNA (10).

A second hypothesis, that may also better explain the aforementioned increases in gene/miRNA expression, are secondary effects. Accordingly, it is possible that the knock-out of miR-122 in clone U6 20.16 caused an increase in expression of direct target genes, some of which may control - positively or negatively - other miRNAs and genes. In fact, we found at least 20 genes that are classified as direct miR-122 targets by miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/index.php) and whose expression is up-regulated in clone U6 20.16 (*e.g.*, TRIB1, LMNB2, FUT8 and others; Supplementary Expression profiles). When factoring in potential off-target miRNAs and genes that may have been inadvertently mutated and dysregulated as well, the complexity of this combinatorial network of secondary effects rapidly escalates and becomes hard to disentangle. Therefore, this mechanism also remains a theoretical possibility that we cannot formally prove or refute at this point.

A third mechanism that we envision is likewise a secondary effect, but on an even more global level, *i.e.*, the entire RNAi machinery of the cell. This is suggested by comprehensive data from us and many others, showing that various steps of the RNAi machinery are rate-limiting and thus can be saturated by high amounts of exogenous RNAi triggers, such as over-expressed shRNAs or miRNAs (*e.g.*, (11-13) and further references therein). Considering that miR-122 is by far the most abundant miRNA in hepatocytes, it is very tempting to speculate that its complete knock-out may free essential components of the RNAi pathway, such as Exportin-5 or Argonaute-2. As a consequence, this could affect processing and/or activity of other miRNAs that likewise depend on this pathway, which could ultimately lead to the observed dysregulation of numerous miRNAs and associated downstream targets (genes/mRNAs).

Last but not least, the perhaps most exciting hypothesis is direct regulation of expression/processing of at least a subset of the perturbed miRNAs in our clone by miR-122. This concept of "miRNA-mediated miRNA regulation" or "miRNA hierarchy system" (14) has only emerged in the last few years, but the evidence is compelling and rapidly increasing. For instance, Tang and colleagues showed that miR-709 is located in the nucleus where it binds to a recognition element in pri-miR-15a/16-1, thereby supervising and preventing its processing into pre-miRNA. Likewise, Matkovich *et al.* reported that transgenic mice which over-express miR-378 or miR-499 in their hearts show dysregulation of up to 30 cardiac miRNAs and up to 1600 mRNAs (15). Of the latter, the majority (about 75%) were classified as secondary hits due to miR-378/-499-mediated dysregulation of other miRNAs, illustrating the remarkable power and complexity of these epitranscriptional networks. Similarly, Chuang *et al.* recently demonstrated that up-regulated miR-494 in hepatocellular carcinoma (HCC) down-regulates various miRNAs that are involved in vascular invasion, including miR-200 and miR-192 (16). An interesting mechanism was proposed whereby miR-494 inhibits the TET1 methylcytosine dioxygenase, which in turn results in CpG

hypermethylation and thus repression of the invasion-suppressive miRNAs, and ultimately HCC. As a final example that is highly relevant to our work, Uhde-Stone *et al.* used TALENs to obtain bi-allelic knock-outs of miR-21 in HEK293 cells (17). Strikingly and highly reminiscent of our own data, they noted a pronounced (over three-fold) dysregulation of 17 out of 760 tested other miRNAs in their cells as well. This substantiates the emerging idea that miRNAs can directly regulate the expression of other miRNAs and thus form complex regulatory circuits of a higher order, *i.e.*, upstream of the effect of miRNAs on target genes. In this respect, it may be rewarding to also investigate our clone U6 20.16 in more detail in the future, including whole genome sequencing and miR-122 rescue experiments, to rule out secondary mutations and to further pinpoint a possible role of miR-122 as master regulator of other cellular miRNAs.

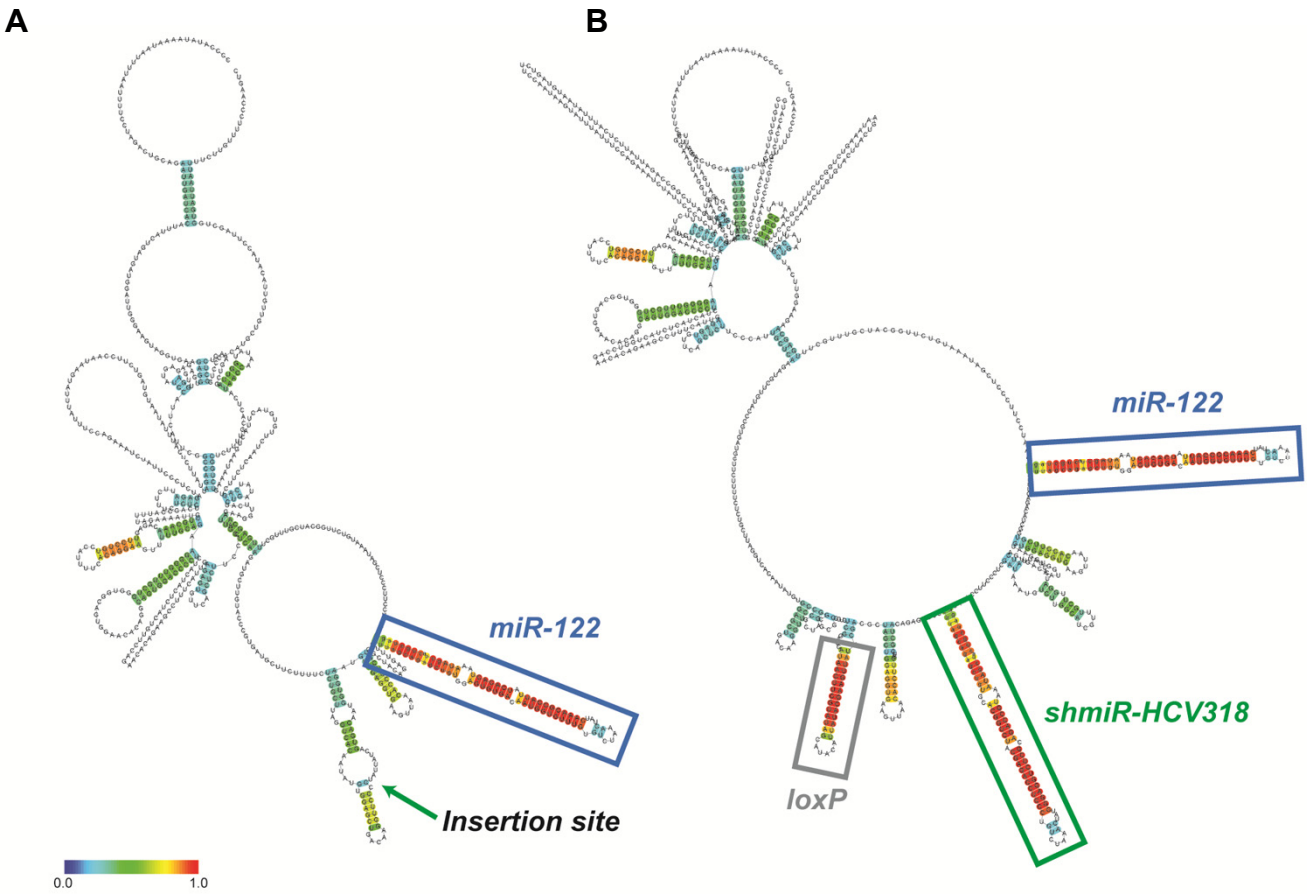
A

CCCCATATAAAATAATTTTAATTTTCCTAGACTGCAGAATTGATCACATTTACTGATGATGGATTGGGAAGTAGGTGAAG
AGGTGAGAGTTGGAGAGTATCCATTCATTCGGCCAGATTATTCTCATTTATAATGTAGTCTTCCAATAAGTATTTATTTC
CAGAAATCTATTCCCTCTAGAGATTCTTTCTCTCTCAGCCATTTTAGAAAATTGTTGCAAACAGAGTTCCTGTCCATTTT
CACAGGAAGTTTTTGCAGAAGGGGTTTGCTGGGTGGCAGTGGAACACAGGCAGTGGACCCCTCTACTACTCTACTGTTCC
AGAACACAGAAGCCTTTCATTGAGAGTGTTCAGCTCTTCCCATTGCTCAAGATGCTTGTACCCGTGATGCTTCTTTTCTC
TGCTTAGGTCACAATATGTGGAGCTGACAAGGTTCCCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACA
CCTTCGTGGCTACAGAGTTTCcttagcagagctgTGGAGTGTGACAATGGTGTTTGTGTCTAAACTATCAAACGCCATTA
TCACACTAAATAgctactgctaggCAATCCTTCCCTCGATAAATGTCTTGGCATCGTTTGCTTTGAGCAAGAAGGTTCAT
CTGATATCAGTCTTCTCAATCTTGTGTACTTACTGAATAAAGTCTGGCTCTTTTGCACTCATAACCAATATGGTTAGCTG
AACCCTGCCTCTCACATGCTGTTGTTACATACCTTAGCTGGTGATTAATTTTCTTGTTTTCCCAAGTC(AATAAAA)

B

CCCCATATAAAATAATTTTAATTTTCCTAGACTGCAGAATTGATCACATTTACTGATGATGGATTGGGAAGTAGGTGAAG
AGGTGAGAGTTGGAGAGTATCCATTCATTCGGCCAGATTATTCTCATTTATAATGTAGTCTTCCAATAAGTATTTATTTC
CAGAAATCTATTCCCTCTAGAGATTCTTTCTCTCTCAGCCATTTTAGAAAATTGTTGCAAACAGAGTTCCTGTCCATTTT
CACAGGAAGTTTTTGCAGAAGGGGTTTGCTGGGTGGCAGTGGAACACAGGCAGTGGACCCCTCTACTACTCTACTGTTCC
AGAACACAGAAGCCTTTCATTGAGAGTGTTCAGCTCTTCCCATTGCTCAAGATGCTTGTACCCGTGATGCTTCTTTTCTC
TGCTTAGGTCACAATATGTGGAGCTGACAAGGTTCCCCGCTAGGGCGCGCCGTCTCGGATC***ATAACTTCGTATAGCATAC***
ATTATACGAAGTTATTCGACGCTAGCTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCcttagcagagctgTGC
ACGGTCTACGAGACCTCCCTGTCTAAACTATGGGAGGTCTCGCAGACCGTAAATAgctactgctaggCAATCCTTCCCTC
GATAAATGTCTTGGCATCGTTTGCTTGGATCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCG
TGGCTACAGAGTTTCcttagcagagctgTGGAGTGTGACAATGGTGTTTGTGTCTAAACTATCAAACGCCATTATCACAC
TAAATAgctactgctaggCAATCCTTCCCTCGATAAATGTCTTGGCATCGTTTGCTTTGAGCAAGAAGGTTCATCTGATA
TCAGTCTTCTCAATCTTGTGTACTTACTGAATAAAGTCTGGCTCTTTTGCACTCATAACCAATATGGTTAGCTGAACCCT
GCCTCTCACATGCTGTTGTTACATACCTTAGCTGGTGATTAATTTTCTTGTTTTCCCAAGTC(AATAAAA)

Supplementary Figure S1. Sequence of the *hcr* locus (hg38: chr18:58450574-58451361) before (A) or after (B) insertion of the anti-HCV shmiRNA. Light green, 5p arm of miR-122; turquoise, 3p arm of miR-122; small letters underlined, flanking 5‘ and 3‘ sequences of pre-miR-122 or or pre-shmiRHCV318, respectively; capital letters underlined, loop sequence; yellow/purple, site of shmiRHCV318 insertion; dark turquoise, 5p arm of shmiRHCV318; ochre, 3p arm of shmiRHCV318; italic/bold, loxP site. The naturally occurring *hcr* poly(A) signal is indicated in brackets.



Supplementary Figure S2. Predicted RNA structure of the 3‘ end of the *hcr* transcript before (A) or after (B) insertion of the anti-HCV shmiRNA (green box in B) 63 nt upstream of the pre-miR-122 hairpin (blue box). The color gradient represents base-pairing probabilities, with red denoting 100% certainty of formation of a specific base pair. Note that shmiRHCV318 insertion does not perturb the pre-miR-122 secondary structure. RNA structure predictions were made using CentroidFold (http://www.ncrna.org/centroidfold/) and using the human genomic sequence hg38: chr18:58450574-58451361 as input (same as in Supplementary Figure S1).

CCCCATATAAAATAATTTTAATTTTCCTAGACTGCAGAATTGATCACATTTACTGATGATGGATTGGGAAGTAGGTGAAG

AGGTGAGAGTTGGAGAGTATCCATTCATTCGGCCAGATTATTCTCATTTATAATGTAGTCTTCCAATAAGTATTTATTTC

CAGAAATCTATTCCCTCTAGAGATTCTTTCTCTCTCAGCCATTTTAGAAAATTGTTGCAAACAGAGTTCCTGTCCATTTT

CACAGGAAGTTTTTGCAGAAGGGGTTTGCTGGGTGGCAGTGGAACACAGGCAGTGGACCCCTCTACTACTCTACTGTTCC

AGAACACAGAAGCCTTTCATTGAGAGTGTTCAGCTCTTCCCATTGCTCAAGATGCTTGTACCCGTGATGCTTCTTTTCTC

TGCTTAGGTCACAATATGTGGAGCTGACAAGGTTCCCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACA

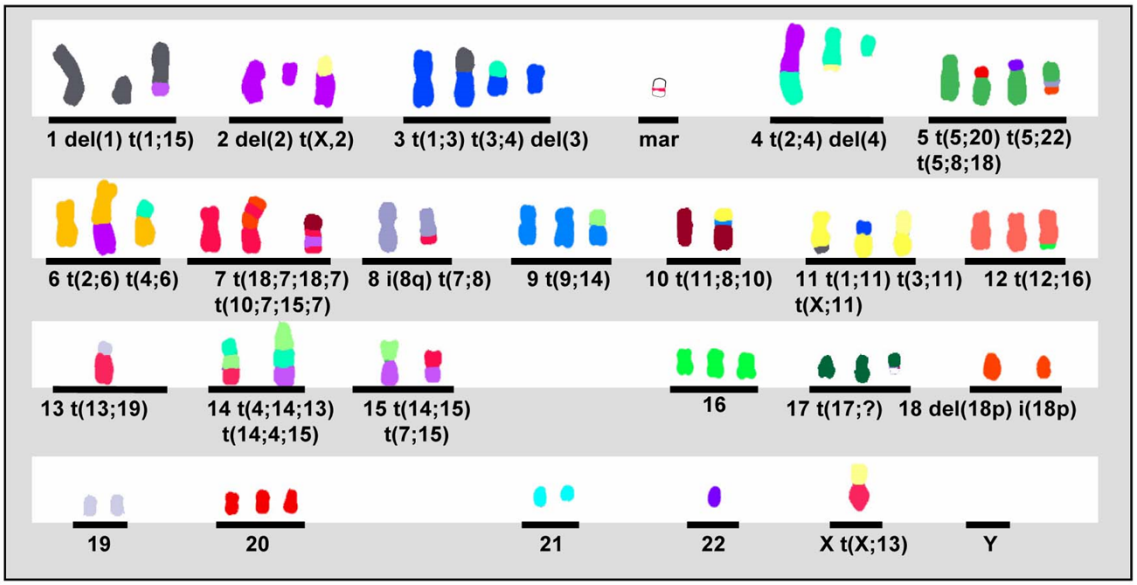
CCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGGAGTGTGACAATGGTGTTTG*TGTCTAAACTATCA*AACGCCATTA

TCACACTAAATAGCTACTGCTAGGCAATCCTTCCCTCGATAAATGTCTTGGCATCGTTTGCTTTGAGCAAGAAGGTTCAT

CTGATATCAGTCTTCTCAATCTTGTGTACTTACTGAATAAAGTCTGGCTCTTTTGCACTCATAACCAATATGGTTAGCTG

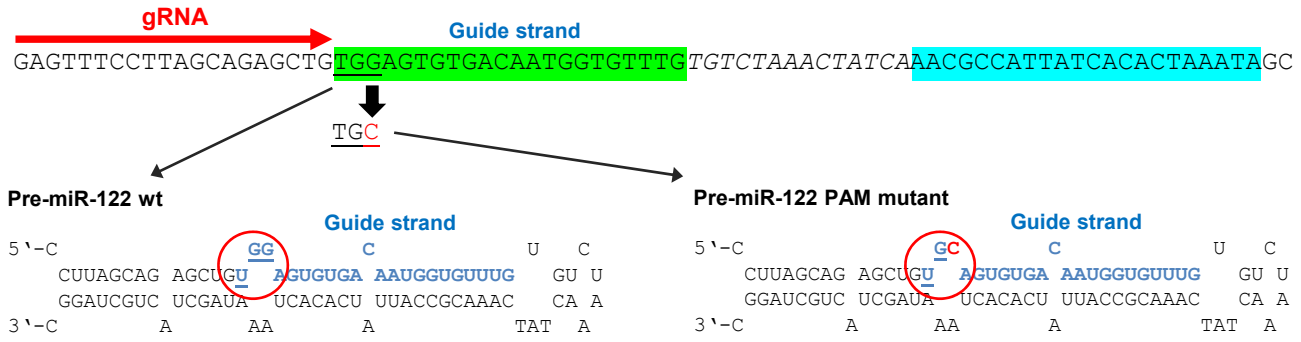
AACCCTGCCTCTCACATGCTGTTGTTACATACCTTAGCTGGTGATTAATTTTCTTGTTTTCCCAAGTC(AATAAAA)

Supplementary Figure S3. Sequence of the *hcr* locus (hg38: chr18:58450574-58451361, same as in Supplementary Figure S1A) with binding sites for our previously published TALEN and CRISPR nucleases (48). The gRNA (red) was driven by the U6 RNA polymerase III promoter. Light green, 5p arm of miR-122; turquoise, 3p arm of miR-122; capital letters underlined, PAM sequence; capital letters in italics, loop sequence.

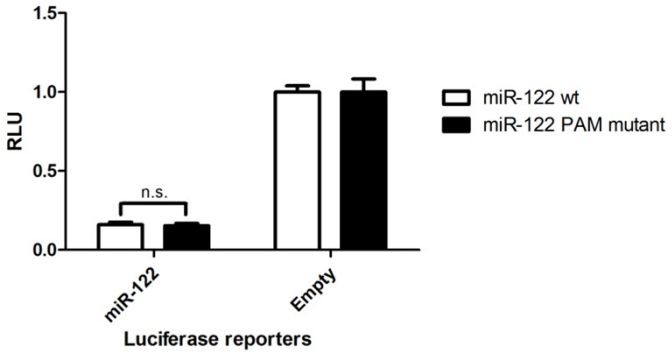


Supplementary Figure S4. M-FISH analysis on chromosomal metaphase spreads of wildtype Huh7 cells. Shown is a 24-color FISH of naïve wildtype Huh7 cells. Chromosomal aberrations - deletions (del), translocations (t) and insertions (i) - as well as the long (q) and short (p) chromosome arms are indicated. mar, marker chromosome.

A



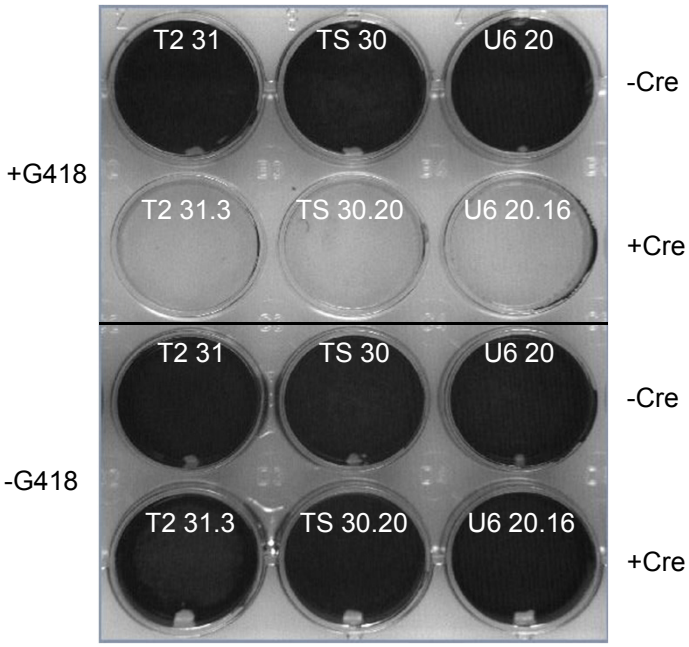
B



C

	Cas9 #1	Cas9 #2	Cas9 #2 + PAM mutant
% HI	3.3%	17.1%	33%
Improvement (ref. Cas9 #1)	-	~5x	10x
Improvement (ref. Cas9 #2)	-	-	~2x

Supplementary Figure S5. Strategies to boost CRISPR-mediated miR-122 editing efficiency. (A) Shown on top is part of the *hcr* locus with the miR-122 hairpin and the binding site for the gRNA used in this study (see also Supplementary Figure S3 for details). Shown underneath is the folded miR-122 hairpin in the wildtype configuration (wt, left) or after mutation of the PAM (right) in the repair template from TGG to TGC (red nucleotide, also circled in red), to prevent adverse cleavage of this template by the gRNA/Cas9. (B) Luciferase assays proving that the PAM mutation does not affect miR-122 functionality. HEK293T cells were co-transfected with the miR-122 expression plasmid from Figure 2B with or without the PAM mutation from panel A, as well as with a miR-122 luciferase reporter (akin to Figure 2C). Statistical analysis was a T-test. n.s., non-significant. (C) Observed increases in homologous integration (HI) efficiency when a more robust Cas9 cDNA (#2) and the repair template with the mutated PAM were used. ref., referred to.



Supplementary Figure S6. Crystal violet staining to confirm loss of G418 resistance after Cre recombinase-mediated excision of the Neo^R cassette. $5x10^4$ cells from the shown clones were seeded per well (24-well plate) and cultured for two weeks with or without G418 as indicated. The cells were then fixed with 4% paraformaldehyde and stained with crystal violet to identify living cells (appearing black). As expected, all cells were killed in the presence of G418 following treatment with the Cre recombinase vectors, confirming efficient elimination of the Neo^R cassette (second row from top).

A

hcr wt allele

```
hcr wt      TCACAATATGTGGAGCTGACAAGGTTCCCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA
T2 31.3     TCACAATATGTGGAGCTGACAAGGTTCCCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [+/-0]
TS 30.20    TCACAATATGTGGAGCTGACAAGGTTCCCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [+1]
U6 6.21     TCACAATATGTGGAGCTGACAAGGTTCCCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAG----GG(miR-122)GCTA [-4]
U6 6.30     TCACAATATGTGGAGCTGACAAGGTTCCCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [+/-0]
U6 7.16     TCACAATATGTGGAGCTGACAAGGTTCCCCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAG-----TGG(miR-122)GCTA [-5]
```

B

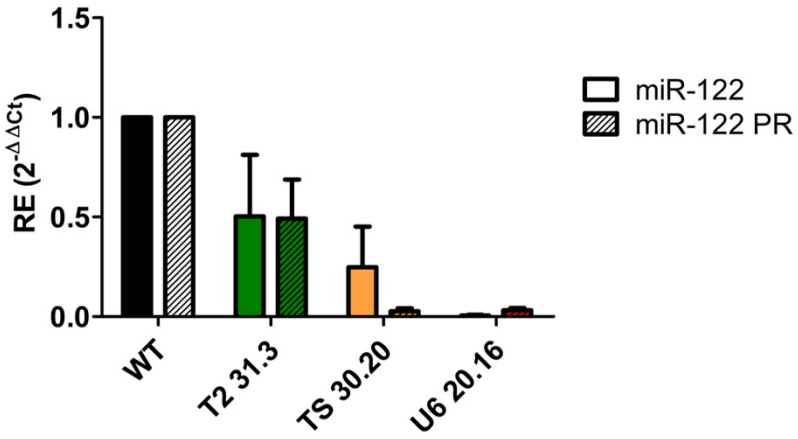
hcr integration allele

```
hcr I       (NeoR_loxP site_shmiRHCV318)CCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA
T2 4.37(H)  (NeoR_loxP site_shmiRHCV318)CCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [+/-0]
T2 31.3          (loxP site_shmiRHCV318)CCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [+/-0]
TS 30.20         (loxP site_shmiRHCV318)CCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [-13]
U6 6.21     (NeoR_loxP site)-------------------[-195n]-------------TGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [-195]
U6 6.30     (NeoR_loxP site_shmiRHCV318)CCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [+/-0]
U6 7.16     (NeoR_loxP site_shmiRHCV318)CCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGC(miR-122)GCTA [+/-0]
U6 20.16(H)      (loxP site_shmiRHCV318)CCTATTATCAGTGACAATGGTGGAATGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGG(miR-122)GCTA [+66]
                                                                                                    [+39]      [+27]
```

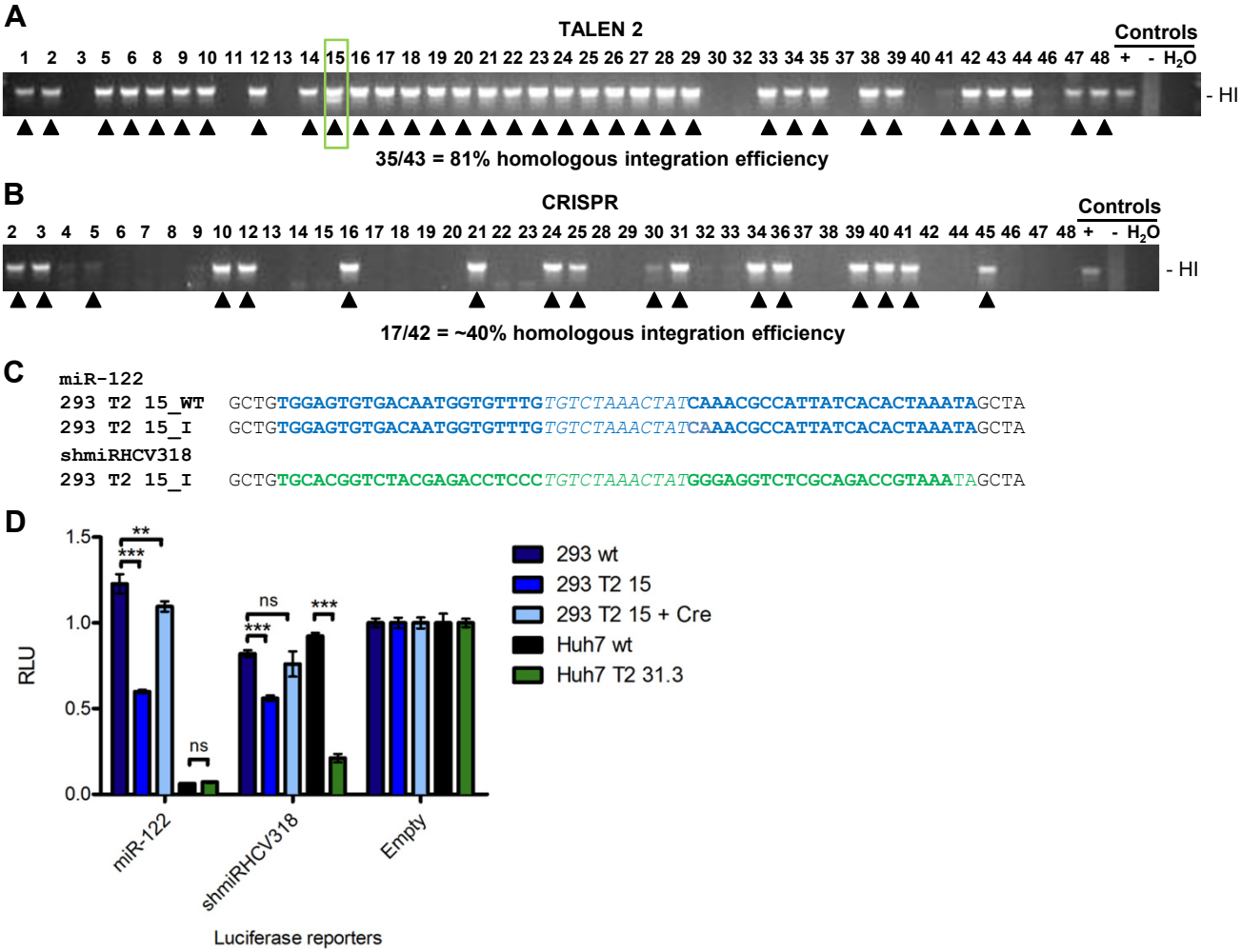
C

```
(miR-122)TGGAGTGTGACAATGGTGTTTGTGTCTAAACTATCAAACGCCATTATCACACTAAATA

(loxP site_shmiRHCV318)ATAACTTCGTATAGCATACATTATACGAAGTTATTCGACGCTAGCTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTGTGCACGGTCTACGAGACCTCCCTGTCTAAACTATGGGAGGTC
TCGCAGACCGTAAATAGCTACTGCTAGGCAATCCTTCCCTCGATAAATGTCTTGGCATCGTTTGCTTGGATCCTATTA

(NeoR)ATGGGATCGGCCATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGT
CAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCC
TGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGAT
GATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGATGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGG
CCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCT
TCTGAGGGGATCAATTCTCTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGT
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA

TS 30.20 [+1] TGGAGTGTGATCAATGGTGTTTGTGTCTAAACTATCAAACGCCATTATCACACTAAATA
                      [+1]

TS 30.20 [-13] TGG-------------TGTTTGTGTCTAAACTATCAAACGCCATTATCACACTAAATA
                    [-13]

U6 20.16 [+66] TGATGAGGGCTTGCTCTTTGCTTTATAGATGGCACTTTCCTGTTGTGATCTGGCATGATGCAAGGGGCAT
                              [+39]                    [+27]
```

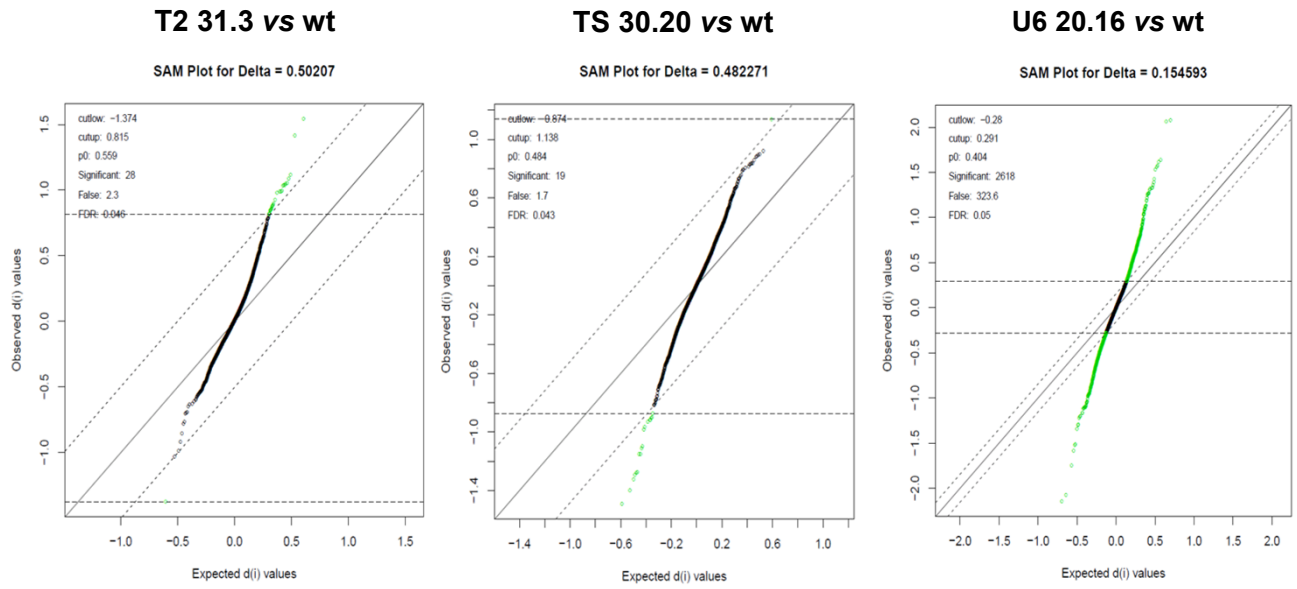
Supplementary Figure S7. Sequencing results of Huh7 cellular clones generated with TALENs or CRISPR/Cas9 and a homologous recombination template containing shmiRHCV318. (A) Sequencing results of *hcr* wildtype (wt) allele. (B) Sequencing results of *hcr* integration (I) allele. The clones in green (T2 31.3), orange (TS 30.20) and red (U6 20.16) were further analyzed in this work as representative examples. Note that for these clones, the shown sequences are already after Cre-mediated excision of the Neo^R cassette. Also note that the G>C mutation (highlighted in red) in the edited allele in clone U6 7.16 results from the use of a repair template with a mutated PAM sequence (see Supplementary Figure S5) and was hence expected. (C) Complete sequences of the elements indicated in parentheses or brackets in panels A/B. In the miR-122 and shmiRHCV318 sequences, the loops are shown in italics. T2, TALEN 2; TS, TALEN Seed; U6, clones generated with U6 promoter-driven gRNA; H, homozygous; TGG, PAM; purple, TALEN 2 binding site (BS); red, gRNA BS; italics in Neo^R cassette, BGH poly(A). In all three panels, numbers of inserted or deleted nucleotides are shown in brackets.



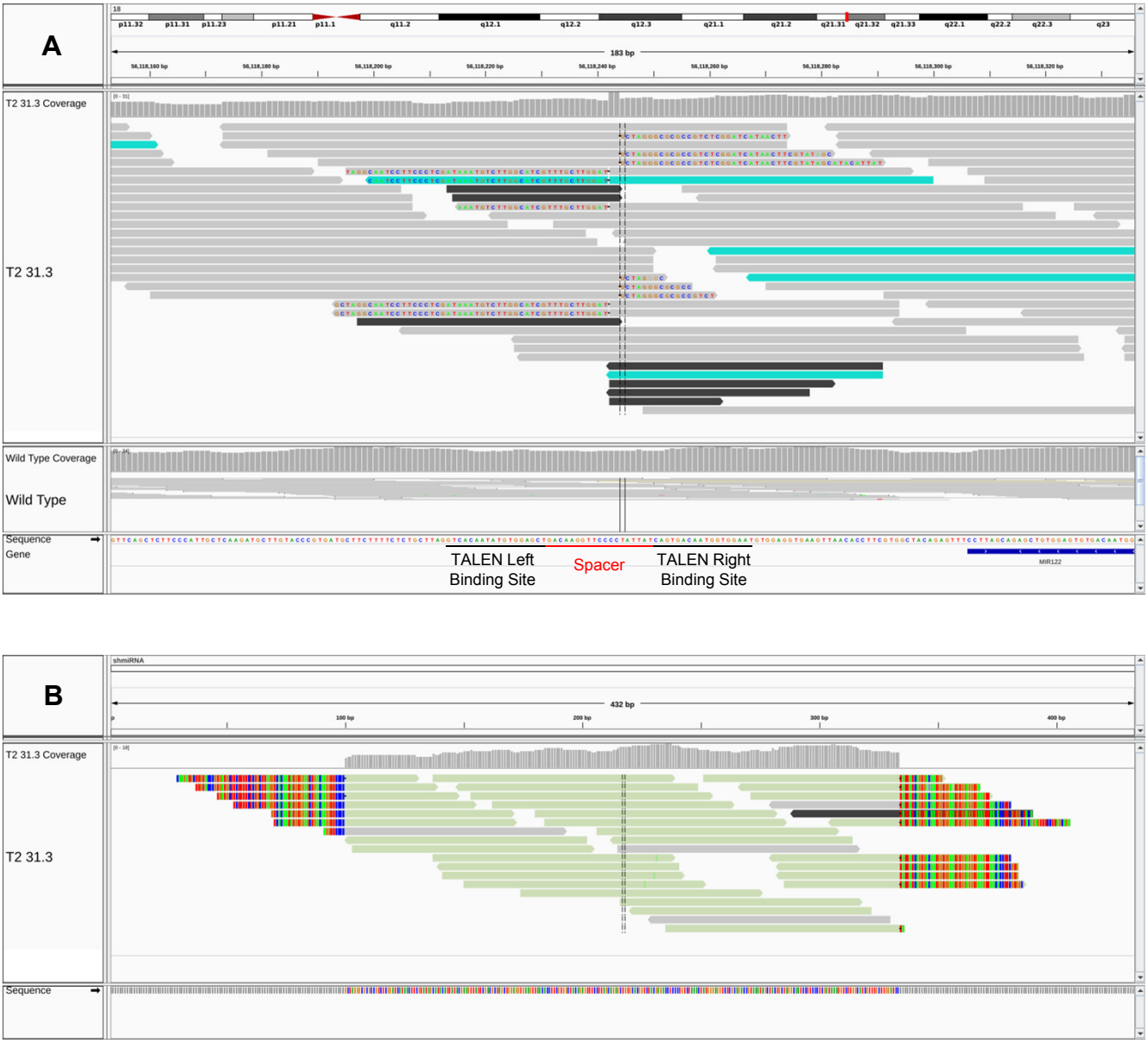
Supplementary Figure S8. qRT-PCR-based quantification of mature miR-122 and miR-122 precursor (PR) steady-state levels. Total RNA harvested at three different cell passages (n=3) was used. Data were normalized (ΔΔCt method) to the small nucleolar RNA SNORD25 and to Huh7 wild-type RNA. Error bars are S.D. Note that no precursor is detected in clone TS 30.20 due to the fact that the 3' end of the forward primer for the precursor is positioned right at the point mutation in miR-122 wild-type allele. Moreover, the deletion in the edited miR-122 allele is also part of the forward primer binding site.



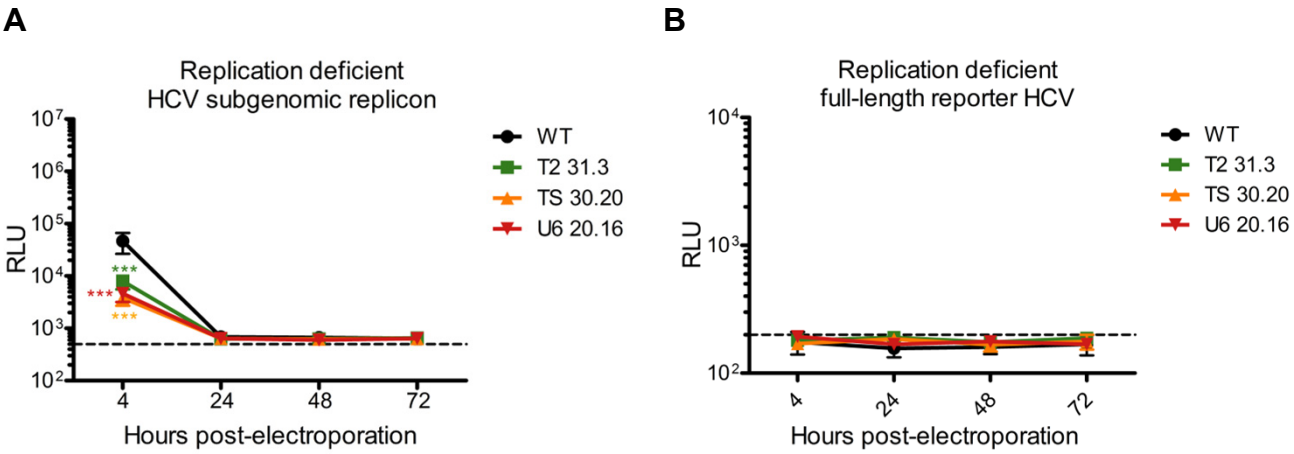
Supplementary Figure S9. Evidence of liver lineage specificity of our strategy. (A/B) PCR showing highly efficient targeting of the *hcr* locus in HEK293 cells using TALEN 2 (A) or CRISPR (B). Arrowheads highlight clones that are positive for homologous integration (HI, yellow primers in Figure 3A). Clone T2 15 (green box) was selected for further analysis. (C) Sequencing results for the wildtype (WT) and engineered (I) alleles of the selected heterozygous HEK293 clone T2 15. Note that the miR-122 hairpin is intact in both alleles. (D) Luciferase assays akin to those in Figure 5 to detect the activity of miR-122 and shmiRHCV318 in HEK293 clone T2 15. Huh7 clone T2 31.3 was included as positive control, together with wildtype Huh7 cells. psiCheck-2 plasmids with perfect binding sites for miR-122 or shmiRHCV318 in the 3‘ UTR of *Renilla* luciferase were used as reporters. Data were normalized to the corresponding Firefly luciferase signal from each plasmid and to psiCheck-2 without miRNA/shmiRNA binding sites (empty, set to 1.0). Statistical analysis was a one-way ANOVA with Dunnett‘s multiple comparison post-test for HEK293 cells and t-test for Huh7 cells. RLU, relative light units; ns, non-significant; ***: p<0.001; **: p<0.01. Error bars are S.D.



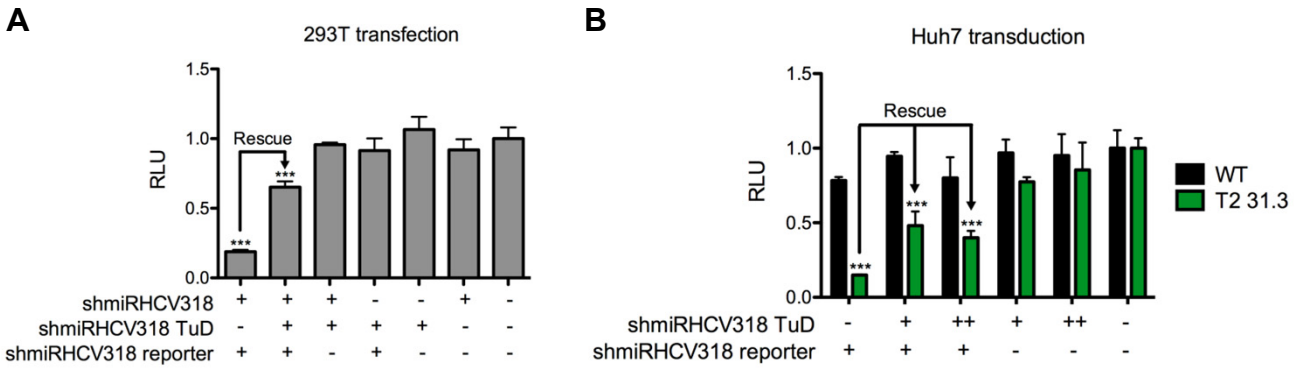
Supplementary Figure S10. QQ plots showing the quantiles of the observed versus the expected statistics for the analysis of differential gene expression between the engineered cellular clones and wildtype (wt) Huh7 cells. See Methods for details. “Delta” is the value of Δ for which the numbers of genes and the estimated FDRs (false discovery rates) are computed. “cutup” and “cutlow” represent the upper and lower cutoffs, respectively, for a gene to be called differentially expressed. “p0” denotes the estimated prior probability that a gene is not differentially expressed. “Significant” is the number of called differentially expressed genes, and “False” the number of falsely called genes. Note that the expected number of false positives is given by p0 x False such that the number of falsely called genes denoted by False is only equal to the expected number of false positives if p0=1. “Called” is the number of genes called differentially expressed. FDR is the estimated FDR computed by p0 x False / FDR.



Supplementary Figure S11. IGV screenshot confirming proper and exclusive integration of the shmiRHCV318 hairpin into the miR-122 locus. (A) Integration site of the shmiRHCV318 hairpin within the miR-122 locus. Soft-clipped regions of the reads are colored and correspond to the start and end of the shmiRNA sequence. (B) Alignments on the shmiRHCV318 contig. Soft- and hard-clipped bases are colored and match the integration site at chr18. Read colors: black, supplemental alignment; grey, read has mate on same contig; light green, mate is on chr18; turquoise, mate is on shmiRHCV318 contig.



Supplementary Figure S12. Negative controls for the experiments in Figure 7. The shown Huh7 cells (wild-type [WT] or engineered) were electroporated with (A) a replication-deficient HCV subgenomic replicon or (B) replication-deficient full-length reporter HCV co-encoding luciferase reporters. Samples were harvested at the shown time points post-electroporation. Depicted data are from three independent experiments (n=3). Dashed lines show background luciferase expressions levels measured with a mock control (no RNA during the electroporation). Statistical analysis is a two-way ANOVA with a Bonferroni's multiple comparison post-test in which every engineered Huh7 clone was compared to Huh7 WT cells. ***: $p<0.001$; no stars: non-significant. Error bars are S.D.



Supplementary Figure S13. Validation of the anti-shmiRHCV318 tough decoy (TuD) used in Figure 7F. (A) HEK293T cells were transfected with the indicated plasmids (+; - were transfections with the corresponding empty plasmids, *i.e.*, not encoding any shmiRNA or TuD, or the original psiCheck-2 luciferase reporter plasmid without shmiRNA binding site). TuD co-expression sequestered and thereby inactivated the shmiRHCV318, in turn partially rescuing expression of the shmiRHCV318-sensitive luciferase reporter (evidenced by the indicated increase in RLUs). (B) Huh7 wildtype or T2 31.3 cells were transduced with AAV-DJ vectors expressing the shmiRHCV318 TuD at $4x10^4$ (+) or $1x10^5$ (++) particles per cell. The lower dose already sufficed to partially rescue the co-transfected shmiRHCV318-sensitive luciferase reporter (evidenced by the indicated increases in RLUs). Data were normalized to the corresponding Firefly luciferase signal from each plasmid and to the empty control (rightmost bars, set to 1.0). n=3. Statistical analysis was a one- or two-way ANOVA with a Bonferroni's multiple comparison post-test. All conditions were compared to the empty controls. ***: $p<0.001$. Error bars are S.D.

Supplementary Table S1. Primers and oligonucleotides used in this study.

Name	Sequence (5'->3')
hcr_5'_1kb_F	aata AGATCT gatat CCTAGG ACTGCAAGTCCAGCTTGAAG
hcr_5'ins1_R1	aata GTCGAC cgat GCTAGC GGGGAACCTTGTCAGCTCCACATATTGTG
hcr_5'ins1_F	taat GCTAGC gcta GGATCC TATTATCAGTGACAATGGTGG
hcr_3'_200bp_R	gcta GTCGAC CAAGATTGAGAAGACTGATATC
shmiR_F1	gtcaa GGTACC TGGAGGTGAAGTTAACACCTTCGTG
shmiR_F2	gtcaa GCTAGC TGGAGGTGAAGTTAACACCTTCGTG
shmiR_R1	cagtt GGATCC AAGCAAACGATGCCAAGACATTTATCG
shmiR_R2	tcga GGATCC GAGACG aca CGTCTC cctag AAGCAAACGATGCCAAGACATTTATC
shmiRHCV318_F	TGTCTGTCTAAACTATGGGAGGTCTCGCAGACCGTAAATAGCTACTGCTAGGC
shmiRHCV318_R	TGTCATAGTTTAGACAGGGAGGTCTCGTAGACCGTGCACAGCTCTGCTAAGG
miRhAAT_F	GCTTTGTCTAAACTATAAGCGTTTAGGAATGTTTACAATAGCTACTGCTAGGC
miRhAAT_R	GCTTATAGTTTAGACAAAGCGTTTAGGCATGTTTAACACAGCTCTGCTAAGG
hcr_1.25kb_5'_F	tgaa CTCGAG TCATGGAAAACTGCCCAGGAGGC
hcr_5'ins1_R2	aata GGATCC cgat GCTAGC GGGGAACCTTGTCAGCTCCACATATTGTG
hcr_1.25kb_3'_R	gata AGATCT GGTTGATCTAACCCTGTCTATTTG
EGFP_tag_F	tact ACTAGT cgctc GCGGCCGC GAGCGCACCATCTTCTTCAAGG
EGFP_tag_R	tact CTCGAG TGTCGCCCTCGAACTTCACCTCG
SV40_p(A)_F	tact ACTAGT CAGACATGATAAGATACATTGATG
SV40_p(A)_R	tactt GCGGCCGC CCACATTTGTAGAGGTTTTACTTGC
MCS_F	ctag GGCGCGCC tat GTCGAC cgcgt GAGACG CGTCTC ttcgac GCTAGC atcgg
MCS_R	gatcccgat GCTAGC gtcgaa GAGACG CGTCTCA cgcg GTCGAC ata GGCGCGCC
loxP_F	cgcgc CGTCTC ggatc ATAACTTCGTATAGCATACATTATACGAAGTTAT
loxP_R	tcga ATAACTTCGTATAATGTATGCTATACGAAGTTAT gatcc GAGACG g
RSV_F	gtta GTCGAC GCTGCTTCGCGATGTACGGGCCAG
RSV_R	cccgggttgtggcaagcttggaggtg CACACCAATG
Neo_F	ctccaagcttgccacaacccggg ATAATTC
BGH_p(A)_R	ttat GGATCC TCCCCAGCATGCCTGCTATTG
hcr5'homology_AvrII	tga CCTAGG TCATGGAAAACTGCCCAGGA
hcr3'homology_SalI	acgc GTCGAC GGTTGATCTAACCCTGTCTA
hcr_ins-check_5	TAGGGTTCGTAGTCATTCTCTCTG
miR-122_F	GAGTTGGAGAGTATCCATTCA
miR-122_F_Not	gagtc GCGGCCGC GAGTTGGAGAGTATCCATTCATTCGGCCAG
miR-122_R_Eco	tgatg GAATTC CAGTCTTAGCCTTCTGCGTCTAACACTTTC
anti-miR-122	CAAACACCATTGTCACACTCCA
shHCV318_sense	TCGATGCACGGTCTACGAGACCTCCC
Neo_RT	CTTGCTCCTGCCGAGAAAGT
Psi_perf_for	tcga GACAAACACCATTGTCACACTCCATCTAGAGC
Psi_perf_rev	ggcc GCTCTAGATGGAGTGTGACAATGGTGTTTGTC
HCV-318_perf_F	tcga GGGAGGTCTCGTAGACCGTGCA
HCV-318_perf_R	ggcc TGCACGGTCTACGAGACCTCCC
hAAT_sense_FWD	tcga GAAGCGTTTAGGCATGTTTAACATC
hAAT_sense_REV	ggcc GATGTTAAACATGCCTAAACGCTTC
shmiR_NheI_F	gtcaa GCTAGC TGGAGGTGAAGTTAACACCTTCGTG
shmiR_BamHI_R	cagtt GGATCC AAGCAAACGATGCCAAGACATTTATCG
TuD_shmiR-HCV318_For	aata CGTCTC gtccca GGGAGGTCTCGTAGACCGTGCACAAGTATTCTGGTCACAGAATA
TuD_shmiR-HCV318_Rev	aata CGTCTC gtcggt TGCACGGTCTACGAGACCTCCCGTTGTATTCTGTGACCAGAATA
tRNAQ1	TGGAGGTTCCACCGAGAT
shmiRHCV318int_rev	AGACTGATATCAGATGAACC
BamHI_Gate_BbsI_F	gatcccaccgagtcttcgcttcact GAAGAC cattttt
HindIII_Gate_BbsI_R	agcttaaaaatggtcttcagtgaagc GAAGAC tcggtg
D1_ccdb_F	cttcctga GAAGAC cacacc GAGTTGGAGAGTATCCATTCATTCGGCCAG
D1_ccdb_R	acgacgtt GAAGAC cgaaaa CAGTCTTAGCCTTCTGCGTCTAACACTTTC
delrepfor_BglII	gatct ATCCTAGTCGTACGCTACGGA

delreprev_HindIII	a TAGGATCAGCATGCGATGCC ttcga
PAM_mut_NGC_hcr_F	ctta GCAGAGCTGTGCAGTGTGACAATGGTG
PAM_mut_NGC_hcr_R	cacc ATTGTCACACTGCACAGCTCTGCTAAG
F_BamHI_Frag1_#752	GTTTGCTTGGATCCTATTATCAGTG
R_KpnI_Frag2_#752	CTGGAGTGCAGTGGTACCATC
F_BamHI_Frag1_#636	GTTCCCCGCTAGCGCTAGGATC
R_AscI_Frag2_#636	CTGATATCGGCGCGCCTGATC

Restriction sites are shown in uppercase and underlined, and overhangs or stuffer sequences are shown in lowercase.

Supplementary Table S2. Private genetic variants in wildtype *versus* engineered Huh7 cells.

	WT	T2 31.3
SNVs only in	12824	8361
Indels only in	2441	1884
SV only in	574	385

Using GRCh37 (version hs37d5) as reference genome, single nucleotide variants (SNVs) and short indels were identified using Platypus (18), and structural variants (SVs) using Manta (19). Only the variants that passed all internal filters within Platypus and Manta were then used in comparisons of wild type Huh7 cells (WT) and our lead clone T2 31.3. In all categories, WT cells show even more private variants than clone T2 31.3. This suggests that the private variants in clone T2 31.3 have accumulated during Huh7 clonal derivation, rather than representing specific nuclease-induced variations.

SUPPLEMENTARY MATERIALS AND METHODS

Plasmids

Plasmid pdsAAV-CMV-hcr (Figure 2B left), expressing a 1.2 kb fragment of the *hcr* locus comprising the miR-122 hairpin (hg38: chr18: 58450012-58451248) from a cytomegalovirus (CMV) promoter, was assembled based on a self-complementary AAV vector construct available in our lab (pBS-sds-decoy-empty) that contains a CMV promoter-driven *gfp* cDNA followed by a 55 bp minimal polyadenylation site (11). Therefore, the *hcr* region was PCR-amplified from genomic Huh7 DNA as two individual fragments, using primers hcr_5'_1kb_F and hcr_5'ins1_R1 (for all primers, see Supplementary Table S1) for fragment #1 (1 kb, ending with NheI and BamHI restriction sites that were introduced by the reverse primer), and primers hcr_5'ins1_F and hcr_3'_200bp_R for fragment #2 (remaining 200 nucleotides, see also Supplementary Figures S1-2 and Figure 2B left). Fragment #1 was double-digested with BglII/SalI and cloned into the BamHI/SalI sites of pBS-sds-decoy-empty. The resulting plasmid intermediate and fragment #2 were both cut with NheI/SalI and ligated together, to replace the *gfp* cDNA in pBS-sds-CMV-hcr-empty and bring the *hcr*/miR-122 region under the control of the CMV promoter, yielding plasmid pdsAAV-CMV-hcr.

In parallel, a construct encoding shmiRHCV318 based on the miR-122 scaffold was created by overlap extension PCR. Therefore, we first amplified 5' and 3' fragments from pcDNA5/FRT/TO-hcr (an expression plasmid available in our lab that encodes the human miR-122 hairpin under a CMV promoter), using primer pairs shmiR_F1 and shmiRHCV318_R, or shmiRHCV318_F and shmiR_R1, respectively. The two fragments were then assembled by overlap extension PCR and cloned into the KpnI/BamHI sites of pcDNA5/FRT/TO (Life Technologies, Darmstadt, Germany). The shmiRHCV318 sequence was next amplified from the resulting plasmid pcDNA5-shmiRHCV318 using primers shmiR_F2 and shmiR_R1, and then inserted into the NheI/BamHI sites of pdsAAV-CMV-hcr, producing plasmid pdsAAV-CMV-hcr/shmiRHCV318 (Figure 2B center top). Likewise, plasmid pdsAAV-CMV-hcr/shmiRhAAT (Figure 2B center bottom) was created by first assembling the shmiRhAAT hairpin (using overlap extension PCR, and primers miRhAAT_F and miRhAAT_R), and by then PCR-amplifying (using primers shmiR_F2 and shmiR_R2) and cloning it into the NheI/BamHI sites in pdsAAV-CMV-hcr. Notably, the reverse primer introduced two inverted BsmBI sites which allowed subsequent insertion of additional hairpins into plasmid pdsAAV-CMV-hcr/shmiRhAAT (purple in Figure 2B). This was exploited for creation of plasmid pdsAAV-CMV-hcr/shmiRhAAT_HCV318 (Figure 2B right). Therefore, shmiRHCV318 was PCR-amplified from pcDNA5-shmiRHCV318 using primers shmiR_F2 and shmiR_R2, cut with NheI/BamHI, and ligated into BsmBI-digested pdsAAV-CMV-hcr/shmiRhAAT. The resulting plasmid pdsAAV-CMV-hcr/shmiRhAAT_HCV318 again contains two inverted BsmBI sites for future insertion of additional hairpins (purple in Figure 2B right).

Luciferase reporters were generated using the psiCheck-2 plasmid (Promega, Mannheim, Germany) as a backbone. A perfect miR-122 binding site was generated by annealing primers Psi_perf_for and

Psi_perf_rev and cloning the resulting product into the XhoI/NotI sites of psiCheck-2. The same approach was used for the shmiRHCV318 or shmiRhAAT perfect binding sites, based on primers HCV-318_perf_F and HCV-318_perf_R, or hAAT_sense_FWD and hAAT_sense_REV, respectively.

The template for homologous recombination of the shmiRHCV318 sequence into the endogenous *hcr* locus of Huh7 cells was generated in a series of cloning steps. We first amplified homology arms of 1.25 kb from genomic DNA of Huh7.5 cells with primers hcr_1.25kb_5'_F and hcr_5'ins1_R2 for the 5' arm, or hcr_5'ins1_F and hcr_1.25kb_3'_R for the 3' arm, respectively. In addition, the SV40 polyadenylation site was amplified from psiCheck-2 using primers SV40_p(A)_F and SV40_p(A)_R, and a small sequence tag corresponding to a portion of the *egfp* coding region was amplified from pBS-U6 (11) using primers EGFP_tag_F and EGFP_tag_R. The SV40 polyadenylation site was digested with SpeI/NotI, the *egfp* tag with NotI/XhoI, and the 5' homology arm with XhoI/BamHI, and all three fragments were inserted in a single ligation reaction into the SpeI/BamHI sites of pSSV9-S2 (a derivative made in our lab from pBlueScript II SK, Agilent Technologies, Santa Clara, CA, USA). Into the BamHI/BglII sites of this cloning intermediate, we next inserted the 3' homology arm (also digested with BamHI/BglII), thereby generating a plasmid containing an SV40 polyadenylation site facing outwards (to shield the vector from possible erroneous transcription upon potential random integration events), and 5' and 3' homology arms separated by NheI and BamHI restriction sites. Next, a multiple cloning site was inserted into the NheI/BamHI sites using annealed oligonucleotides MCS_F and MCS_R. The resulting plasmid was cut with BsmBI, and a Cre recombinase recognition site (loxP) was inserted as annealed oligonucleotides loxP_F and loxP_R (carrying AscI/SalI overhangs compatible to those from the BsmBI digest). This plasmid was digested with SalI/BsmBI (effectively creating SalI/BamHI overhangs), to accommodate a RSV promoter-driven Neo^R cassette comprising a bovine growth hormone polyadenylation site (Neo-bgh poly[A]). The latter was generated by PCR-amplifying (i) the RSV promoter from pBS-U6 using primers RSV_F and RSV_R, and (ii) Neo-bgh poly(A) from pIRESNeo-hAgo2-GA (20) with primers Neo_F and BGH_p(A)_R, followed by assembly via overlap extension PCR. It was then also digested with SalI/BsmBI for cloning into the loxP-carrying plasmid described above. A second loxP site was introduced into AscI/SalI sites upstream of the Neo^R cassette, by inserting annealed oligonucleotides loxP_F and loxP_R, generating pSSV9-hcr-donor-empty. Finally, the shmiRHCV318 hairpin was amplified from plasmid pdsAAV-CMV-hcr/miRHCV318 using primers shmiR_NheI_F and shmiR_BamHI_R, and cloned into the NheI/BamHI sites of pSSV9-hcr-donor-empty, placing it right upstream of the 3' *hcr* homology arm and creating the targeting vector pSSV9-hcr-donor-shmiRHCV318.

The anti-miR-122 TALEN and CRISPR/Cas9 expression plasmids have recently been described by us (10). The vector for expression of *Streptococcus pyogenes* Cas9 used in all initial experiments and called "Cas9 #1" in Supplementary Figure S5C was originally published by the Church lab (21) and also used in our own prior publication (10). The "Cas9 #2" variant which gave roughly five-fold higher miR-122 editing efficiencies (Supplementary Figure S5C) was generated by replacing the Cas9 cDNA with another codon-optimized variant published by the Zhang lab (22). Details of the generation of this construct will be

reported elsewhere (Schmidt *et al*., manuscript in preparation). Note that the use of the less efficient "Cas9 #1" variant also explains the moderate frequencies of up to around 20% miR-122 editing observed in our prior report (10), which under-estimates the actual potency of the anti-miR-122 gRNA.

We have previously published AAV vector constructs expressing Cre recombinase or YFP (yellow fluorescence protein) under a liver-specific transthyretin (TTR) promoter (23). To generate the AAV-CAG-Cre:GFP vector, the single-stranded AAV backbone pSSV9-S2 (see above) was used. First, the AAV2 *rep* sequence was eliminated from pSSV9-S2 using annealed primers delrepfor_BglII and delreprev_HindIII. Then, the CAG-Cre:GFP cassette was obtained from plasmid pCAG-Cre:GFP (plasmid #133776, Addgene, Cambridge, MA, USA) via SpeI/HindIII digestion and subsequently subcloned into the SpeI/HindIII-digested and *rep*-depleted pSSV9 backbone generated in the previous step.

The constructs used for the HCV replication inhibition experiments (full-length HCV reporter virus JCR2a, as well as subgenomic HCV replicon JFH1 and a Dengue virus replicon) were all reported before (24-26).

To create the shmiRHCV318 tough decoy (TuD) used in Figure 7F and Supplementary Figure S13, primers TuD_shmiR-HCV318_For and TuD_shmiR-HCV318_Rev were used for overlap extension PCR. The PCR product was then digested with BsmBI and cloned into the AAV vector construct pBS-U6-TuD-empty (27).

For sequencing of miR-122-edited clones shown in Supplementary Figure S7, a vector comprising the negative selection marker *ccdB* flanked by BbsI restriction sites was generated. To this end, oligonucleotides BamHI_Gate_BbsI_F and HindIII_Gate_BbsI_R were annealed and cloned into pBluescript KS(-) (Agilent Technologies) via BamHI/HindIII. Next, the *ccdB* open reading frame was obtained by BbsI digest from our U6-shRNA expression vectors (Schmidt *et al*., manuscript in preparation) and ligated into the modified pBluescript plasmid using the BbsI sites introduced by the oligonucleotide modification in the previous cloning step. The genomic locus of single homologous integration clones was amplified using primers D1_ccdb_F and D1_ccdb_R. PCR products were purified and cloned via Golden Gate assembly using 40 fmol of each, PCR product and backbone. The final reaction contained 1 µl of T4 DNA Ligase, 0.75 µl BbsI, 1 µl buffer 2.1 (all NEB, Ipswich, MA, USA) as well as 1 mM of both, ATP and DTT in 20 µl total volume. The Golden Gate Assembly was conducted in a thermocycler using 10 cycles of 37°C 3 minutes, 16°C 5 minutes, followed by a final cycle of 16°C for 15 minutes and heat inactivation (75°C, 20 minutes). Finally, 10 µl of this reaction were transformed into 50 µl of chemo-competent *E. coli*. Single bacterial clones were picked, plasmid DNA isolated and sent for sequencing using Primer shmiRHCV318int_rev.

To mutate the PAM in our homologous recombination template and in the miR-122 expression vector described above, the desired GG-to-GC mutation was introduced via overlap-extension PCR using primers PAM_mut_NGC_hcr_F and PAM_mut_NGC_hcr_R. In more detail, two fragments were amplified from the homologous recombination template (pSSV9-hcr-donor-shmiRHCV318) using primers F_BamHI_Frag1_#752 and PAM_mut_NGC_hcr_R, as well as PAM_mut_NGC_hcr_F and

R_KpnI_Frag2_#752, respectively. For the miR-122 expression vector (pdsAAV-CMV-hcr), primers F_BamHI_Frag1_#636 and R_AscI_Frag2_#636 were used instead. PCR fragments 1 and 2 were fused in a second PCR again using primers F_BamHI_Frag1_#752 and R_KpnI_Frag2_#752, or F_BamHI_Frag1_#636 and R_AscI_Frag2_#636, respectively. PCR products were finally cloned via BamHI/KpnI or BamHI/AscI into pSSV9-hcr-donor-shmiRHCV318 or pdsAAV-CMV-hcr, respectively.

Cell culture and AAV vector production

Huh7 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Life Technologies) supplemented with 10% fetal bovine serum (Sigma, Steinheim, Germany), 100 U/ml penicillin and 100 µg/ml streptomycin (Life Technologies), as well as 1% non-essential amino acids (Life Technologies). Huh7 clones which still contained the floxed Neo^R cassette (before treatment with Cre recombinase) were cultured in Huh7 media supplemented with 500 µg/ml G418 (Life Technologies). After treatment with Cre recombinase, the cellular clones were again cultured in regular Huh7 media without G418. HEK293T cells were cultured in the same complete DMEM used for Huh7, but without non-essential amino acids.

For AAV-DJ CAG-Cre vector production, $4.2x10^6$ HEK293T cells were seeded per dish in 10 15 cm dishes (Thermo Scientific, Darmstadt, Germany). After 48 hours, the cells were triple-transfected using polyethylenimine (PEI) (Polysciences, Eppelheim, Germany). A total of 44 µg DNA was used per plate, comprising 14.6 µg of an adenoviral plasmid providing helper functions for AAV production (28), 14.6 µg of a helper plasmid encoding the chimeric AAV capsid DJ (29) together with AAV-2 Rep proteins, and 14.6 µg of a self-complementary AAV vector genome consisting of two AAV inverted terminal repeats (ITRs) flanking a CAG promoter-driven Cre recombinase (23). The amount of each DNA necessary for 10 dishes was diluted in H_2O to a total volume of 7.9 ml and mixed with the same volume of 300 mM NaCl. Next, 3.52 ml of PEI were diluted in 4.38 ml of H_2O and mixed with 7.9 ml of 300 mM NaCl. The PEI mix was added dropwise to the DNA mix followed by brief vortexing. The mix was incubated at room temperature for 10 minutes, before 3.2 ml were added dropwise to each of the 10 dishes. Cells were harvested 72 hours post-transfection, and AAV vectors were purified using iodixanol gradient ultracentrifugation (OptiPrep, Axis-shield, Dundee, UK) (30).

The self-complementary AAV-DJ vector encoding a U6 promoter-driven TuD against shmiRHCV318 was produced via the same triple-transfection procedure but using 30 instead of 10 dishes. This vector was purified using cesium chloride (Sigma-Aldrich, St. Louis, MO, USA) density gradient ultracentrifugation, followed by dialysis (Slide-A-Lyzer G2 Dialysis Cassettes, Life Technologies) and concentration (Amicon Ultra-15 Centrifugal Filter Units, Merck Millipore, Schwalbach, Germany).

Transfections

For the shmiRNA integration experiments, $5x10^4$ Huh7 or HEK293 wildtype cells per well were seeded in 24-well plates. Twenty-four hours after seeding, the cells were transfected using Lipofectamine 2000 (Life Technologies) following the manufacturer's instructions. Briefly, 800 ng of DNA and 2 µl of Lipofectamine

2000 were used per well. The DNA was diluted in 50 µl of DMEM without any supplements. Of the 800 ng of DNA, 200 ng corresponded to the TALEN (100 ng of each subunit) or the CRISPR/Cas9 (100 ng of Cas9 and 100 ng of the gRNA) expression plasmids (10), while the remaining 600 ng corresponded to the template plasmid for homologous recombination. The media was changed 5 hours after transfection to complete DMEM (see above). The cells were harvested 48 to 72 hours post-transfection, transferred to 15 cm dishes and cultured under G418 pressure (see Methods).

To verify the possibility to co-express shmiRNA together with miR-122 from the *hcr* locus, $3x10^6$ HEK293T cells were seeded in 96-well plates and transfected with the constructs shown in Figure 2B and their corresponding luciferase reporters. Twenty-four hours after seeding, the cells were transfected using Lipofectamine 2000 following the manufacturer's instructions. Briefly, 200 ng of DNA and 0.5 µl of Lipofectamine 2000 were used per well. The DNA was diluted in 25 µl of DMEM without any supplements. Of the 200 ng of DNA, 10 ng corresponded to the respective luciferase reporters (psiCheck-2 with binding sites for miR-122, shmiRHCV318, hAAT or none), 20 ng corresponded to the expression constructs and the remaining 170 ng were stuffer DNA (pIRESneo-FLAG/HA-YFP (31)). All transfections were performed in triplicates. The cells were harvested 48 to 72 hours post-transfection.

For luciferase reporter assays to determine expression and functionality of miR-122 and shmiRHCV318, $1.2x10^4$ wildtype or engineered Huh7 cells per well were seeded in 96-well plates. Twenty-four hours after seeding, the cells were transfected using Lipofectamine 2000 as described above. Of the 200 ng of DNA, 10 ng corresponded to the respective luciferase reporters (psiCheck-2 with binding sites for miR-122 or shmiRHCV318, or no binding sites), while the remaining 190 ng were stuffer DNA. All transfections were performed in triplicates. The media was changed 5 hours after transfection to complete media for Huh7 cells (see above) with or without G418, depending on the cell line. The cells were harvested 48 to 72 hours post-transfection.

PCR reactions

To detect successful homologous recombination, genomic DNA was extracted from the clones grown after transfection of the different nuclease and recombination template combinations and analyzed by PCR. Therefore, a forward primer was used that bound in the *hcr* locus right upstream of the region corresponding to the 5' homology arm in the homologous recombination template (primer hcr_ins-check 5, forward yellow arrow in Figure 3A). The reverse primer bound in the RSV promoter driving the expression of the Neo^R cassette (primer RSV_R, reverse yellow arrow in Figure 3A). A touchdown PCR was performed using Phusion Hotstart II polymerase (Thermo Scientific) and Phusion HF buffer following the manufacturer's instructions. PCR cycling conditions were: initial denaturation at 98°C for 30 seconds; 12 cycles of denaturation at 98°C for 45 seconds followed by primer annealing at 64°C (the temperature was reduced by 0.5°C in each cycle) for 30 seconds and extension at 72°C for 45 seconds; 28 cycles of denaturation at 98°C for 30 seconds followed by primer annealing at 58°C for 30 seconds and extension

at 72°C for 45 seconds; final extension round at 72°C for 10 minutes. The expected size of the PCR product was approximately 1.3 kb (Figure 3B).

To detect minicircle DNA formation after Cre recombinase-mediated excision of the Neo^R cassette, a PCR was performed on genomic DNA extracted from the pool of cells 48 to 72 hours after AAV-Cre recombinase transduction. Primer RSV_R was used as reverse primer, and primer Neo_RT (purple arrows in Figure 3D) as forward primer. This primer combination could amplify the Neo^R cassette only if it had been excised and formed a minicircle. The PCR was performed using Phusion Hotstart II polymerase and Phusion HF buffer following the manufacturer's instructions. PCR cycling conditions were: initial denaturation at 98°C for 30 seconds; 35 cycles of denaturation at 98°C for 10 seconds followed by primer annealing at 57°C for 30 seconds and extension at 72°C for 30 seconds; final extension round at 72°C for 8 minutes. The expected size of the PCR product was 762 nucleotides (Figure 3E top).

To verify proper excision of the Neo^R cassette after Cre recombinase treatment, and to assess whether the integration was hetero- or homozygous, genomic DNA from the colonies was analyzed by PCR using primers hcr5'homology_AvrII (forward) and hcr3'homology_SalI (reverse, red arrows in Figure 3D). The PCR was performed using Phusion Hotstart II polymerase, Phusion GC buffer and DMSO (Thermo Scientific) following the manufacturer's instructions. PCR cycling conditions were: initial denaturation at 98°C for 30 seconds; 30 cycles of denaturation at 98°C for 15 seconds followed by primer annealing at 69°C for 30 seconds and extension at 72°C for 90 seconds; final extension round at 72°C for 10 minutes. The expected size of the PCR products were (all approximate numbers) 2.5 kb for the wildtype amplicon, 2.7 kb for the amplicon after homologous integration and after excision of the Neo^R cassette, and 4.3 kb for the amplicon after homologous recombination but before excision of the Neo^R cassette (Figure 3E bottom).

To sequence the engineered *hcr* locus (Supplementary Figure S7), it was PCR-amplified using primers miR-122_F_Not and miR-122_R_Eco. The PCR was performed using Phusion Hotstart II polymerase, Phusion HF buffer and DMSO following the manufacturer's instructions. PCR cycling conditions were: initial denaturation at 98°C for 30 seconds; 30 cycles of denaturation at 98°C for 10 seconds followed by primer annealing at 57°C for 30 seconds and extension at 72°C for 30 seconds; final extension round at 72°C for 8 minutes. The expected sizes of the PCR products were (all approximate numbers) 1.1 kb for the wildtype amplicon and 1.3 kb for the amplicon with the integrated shmiRNA. The products were then digested with NotI and EcoRI (New England BioLabs, Frankfurt, Germany) and subcloned into NotI/EcoRI-cut pIRESneo-FLAG/HA EYFP. Positive ligation products were sequenced using the GATC LIGHT*run* 96 service (GATC, Konstanz, Germany) and primers miR-122_F or shmiRHCV318int_rev.

Overlap extension PCR was performed using Phusion Hotstart II polymerase and Phusion HF buffer following the manufacturer's instructions and 1 μl of each primer at 100 μM. PCR cycling conditions were: initial denaturation at 98°C for 30 seconds; 30 cycles of denaturation at 98°C for 10 seconds followed by

primer annealing at 55°C for 10 seconds and extension at 72°C for 10 seconds; final extension round at 72°C for 1 minute.

Luciferase reporter assays

Luciferase reporter assays were performed using the Dual-Luciferase Reporter Assay System (Promega) following the manufacturer's instructions. Briefly, the media was removed from the cells 48 hours post-transfection before the cells were washed with 1x PBS. Next, 30 µl of Passive Lysis Buffer were added per well, and the cells were incubated on a rocker platform for 15 minutes at room temperature. Five µl of each lysate were transferred to a white Lumitrac 200 plate (Greiner Bio-One, Frickenhausen, Germany) for measurement in a Glomax 96 Microplate Luminometer (Promega). *Renilla* and Firefly luciferase activities were quantified after consecutive injection into each well of 30 µl of reconstituted Luciferase Assay Buffer and 30 µl of Stop&Glow Buffer supplemented with *Renilla* luciferase substrate.

Quantitative real-time (qRT-)PCR

To quantify the relative expression of miR-122 and shmiRHCV318, quantitative real-time PCR (qRT-PCR) was performed using the miScript system from Qiagen (Hilden, Germany). Cells were harvested from one confluent 10 cm dish and resuspended in 5 ml of QIAzol Lysis Reagent (Qiagen), for purification of total RNA using the provided QIAzol RNA extraction protocol.

For cDNA synthesis, 500 ng of each RNA diluted in RNase-free H_2O to a final volume of 6 µl were reverse transcribed using (per sample) 1 µl of miScript Reverse Transcriptase Mix, 2 µl of 5x miScript HiFlex Buffer and 1 µl of 10x miScript Nucleics Mix. For the control without reverse transcriptase, RNA from wildtype Huh7 cells was used. The reverse transcription mix was incubated at 37°C for 60 minutes, followed by a 5 minute incubation at 95°C to inactivate the miScript Reverse Transcriptase Mix. The cDNAs were diluted 1:10 in RNase-free H_2O and stored at -20°C.

The standard curve to quantify the number of shmiRHCV318 molecules was performed following Qiagen's "Protocol for use of miScript miRNA mimic as a positive control in miRNA detection" (Appendix C, miScript PCR System Handbook). Briefly, the custom designed miScript shmiRHCV318 mimic (MIM0360326, Qiagen) was diluted to a concentration of $1x10^{10}$ copies/µl (stock at 20 µM, size 44 nt) and mixed with 50 ng of carrier RNA (total RNA from wildtype Huh7 cells). This mix was reverse transcribed using 2 µl of miScript Reverse Transcriptase Mix, 4 µl of 5x miScript HiFlex Buffer and 2 µl of 10x miScript Nucleics Mix and incubated as described above. The cDNA was next diluted in 480 µl of carrier RNA (at a concentration of 1 ng/µl), to yield a DNA molecule copy number of $1x10^8$, assuming a 100% efficiency of the reaction. Six serial 1:10 dilutions of the cDNA were then prepared in carrier RNA (1 ng/µl) and stored at -20°C.

The qRT-PCRs were performed using the miScript SYBR Green PCR Kit (Qiagen) and the Hs_miR-122a_1 miScript Primer Assay (MS00003416, Qiagen) to quantify relative miR-122 expression, , the Hs_mir-122_PR_1 miScript Precursor Assay (MP00000308, Qiagen) to quantify relative miR-122

precursor expression, the shmiRHCV318 miScript Custom Primer Assay (MS0074622, Qiagen) to quantify relative shmiRHCV318 expression, and the Hs_SNORD25_11 (MS00014007, Qiagen) control to normalize the expression of miR-122, miR-122 precursor and shmiRHCV318. Per sample, the reaction mix consisted of 7.5 µl of QuantiTect SYBR Green PCR Master Mix, 1.5 µl of 10x miScript Universal Primer, 1.5 µl of the specific 10x miScript Primer Assay and 2.5 µl of RNase-free H_2O. Using a QIAgility pipetting robot, 2 µl of the cDNA template (samples or standard) as well as 13 µl of the PCR reaction mix were transferred to a Qiagen Rotor-Disc. The reactions were run in a Qiagen Rotor-Gene 6000 using the following program: 95°C for 15 minutes followed by 50 cycles of 94°C for 15 seconds, 55°C for 35 seconds and 68°C for 30 seconds. A melting curve was calculated in a last step in which the temperature was increased from 55°C to 95°C in 5-second steps of 1°C each. Data analysis was performed using the Rotor-Gene Q Series Software (Qiagen). The number of shmiRHCV318 copies per cell was determined by multiplying the calculated concentration (based on the standard curve) with the cDNA and RNA dilutions, and by dividing this product by the number of cells counted on the day of harvest ($x = \frac{Calc.\ conc.\ x\ cDNA\ dilution\ x\ RNA\ dilution)}{Number\ of\ cells}$).

Small-RNA Northern blotting

Total RNA was prepared using QIAzol (Qiagen) according to the manufacturer's instructions. Eight µg of total RNA were precipitated, washed once with 70% ethanol and dissolved in 5 µl RNA loading buffer (98% formamide, 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol), briefly denatured at 70°C and resolved in a 15% acrylamide / 8 M urea gel. The gel was electroblotted onto a Nylon membrane (Hybond N+, Amersham, Freiburg, Germany), which was then UV-crosslinked (energy program, HL-2000 Hybrilinker, UVP) and hybridized in ExpressHyb hybridization buffer (Takare/Clontech, Saint-Germain-en-Laye, France) with radioactively labeled probes. Radioactive end-labeling of oligonucleotide probes was performed using T4 polynucleotide kinase (New England BioLabs) under standard reaction conditions. The oligonucleotides used as probes were anti-human transfer RNA glutamine 1 (tRNAQ1) as housekeeper, anti-miR-122 and shHCV318_sense (Supplementary Table S1).

Multiplex Fluorescence-*In-Situ*-Hybridization (M-FISH)

MetaSystems XCyting DNA Probe Kit (MetaSystems GmbH, Altlussheim, Germany) was used to visualize at least 24 chromosomes (22 autosomes plus XY) according to the manufacturer's recommendations. Briefly, low-passage Huh7 cells were cultured in complete DMEM medium at 60% cell confluency, followed by metaphase chromosome preparation on glass slides. After consecutive incubation with 100%, 90% and 70% ethanol, remaining cytoplasmic proteins were removed by pepsin digests. After an additional ethanol series with 100%, 70%, 50% and 30% ethanol, chromosomes were denatured (0.07 N NaOH) and incubated with denatured probes (75°C for 30 minutes) at 37°C for 30 minutes. After hybridization, unbound probes were washed off with 1xSSC buffer and chromosomes counterstained with DAPI, followed by a 1xPBS wash step prior to overlaying with a coverslip. The

detection of at least 24 different chromosome painting probes was achieved with five varicolored fluorochromes by fluorescence microscopy analysis. Multiple metaphase spreads were evaluated, of which a representative picture is shown in Supplementary Figure S4.

Statistical analysis

The statistical tests used were either a one- or two-way ANOVA, depending on the number of variables, with a Bonferroni's multiple comparison post-test in which every engineered Huh7 clone was compared to Huh7 wildtype cells. Statistical analysis was performed using GraphPad Prism software (GraphPad Software, Inc., La Jolla, CA, USA).

SUPPLEMENTARY REFERENCES

1. Conrad, T., Marsico, A., Gehre, M. and Orom, U.A. (2014) Microprocessor activity controls differential miRNA biogenesis In Vivo. *Cell Rep.*, **9**, 542-554.
2. Ballarino, M., Pagano, F., Girardi, E., Morlando, M., Cacchiarelli, D., Marchioni, M., Proudfoot, N.J. and Bozzoni, I. (2009) Coupled RNA processing and transcription of intergenic primary microRNAs. *Mol. Cell. Biol.*, **29**, 5632-5638.
3. Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I. and Proudfoot, N.J. (2008) Primary microRNA transcripts are processed co-transcriptionally. *Nat. Struct. Mol. Biol.*, **15**, 902-909.
4. Chaulk, S.G., Thede, G.L., Kent, O.A., Xu, Z., Gesner, E.M., Veldhoen, R.A., Khanna, S.K., Goping, I.S., MacMillan, A.M., Mendell, J.T. *et al.* (2011) Role of pri-miRNA tertiary structure in miR-17~92 miRNA biogenesis. *RNA Biol.*, **8**, 1105-1114.
5. Yang, X., Marcucci, K., Anguela, X. and Couto, L.B. (2013) Preclinical evaluation of an anti-HCV miRNA cluster for treatment of HCV infection. *Mol. Ther.*, **21**, 588-601.
6. Chakraborty, S., Mehtab, S., Patwardhan, A. and Krishnan, Y. (2012) Pri-miR-17-92a transcript folds into a tertiary structure and autoregulates its processing. *RNA*, **18**, 1014-1028.
7. Haar, J., Contrant, M., Bernhardt, K., Feederle, R., Diederichs, S., Pfeffer, S. and Delecluse, H.J. (2016) The expression of a viral microRNA is regulated by clustering to allow optimal B cell transformation. *Nucleic Acids Res.*, **44**, 1326-1341.
8. Merino, E.J., Wilkinson, K.A., Coughlan, J.L. and Weeks, K.M. (2005) RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.*, **127**, 4223-4231.
9. Wilkinson, K.A., Merino, E.J. and Weeks, K.M. (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.*, **1**, 1610-1616.
10. Senis, E., Fatouros, C., Grosse, S., Wiedtke, E., Niopek, D., Mueller, A.K., Borner, K. and Grimm, D. (2014) CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. *Biotechnol. J.*, **9**, 1402-1412.
11. Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F. and Kay, M.A. (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature*, **441**, 537-541.
12. Khan, A.A., Betel, D., Miller, M.L., Sander, C., Leslie, C.S. and Marks, D.S. (2009) Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. *Nat. Biotechnol.*, **27**, 549-555.
13. Grimm, D. (2011) The dose can make the poison: lessons learned from adverse in vivo toxicities caused by RNAi overexpression. *Silence*, **2**, 8.
14. Tang, R., Li, L., Zhu, D., Hou, D., Cao, T., Gu, H., Zhang, J., Chen, J., Zhang, C.Y. and Zen, K. (2012) Mouse miRNA-709 directly regulates miRNA-15a/16-1 biogenesis at the

posttranscriptional level in the nucleus: evidence for a microRNA hierarchy system. *Cell Res.*, **22**, 504-515.
15. Matkovich, S.J., Hu, Y. and Dorn, G.W., 2nd. (2013) Regulation of cardiac microRNAs by cardiac microRNAs. *Circ. Res.*, **113**, 62-71.
16. Chuang, K.H., Whitney-Miller, C.L., Chu, C.Y., Zhou, Z., Dokus, M.K., Schmit, S. and Barry, C.T. (2015) MicroRNA-494 is a master epigenetic regulator of multiple invasion-suppressor microRNAs by targeting ten eleven translocation 1 in invasive human hepatocellular carcinoma tumors. *Hepatology*, **62**, 466-480.
17. Uhde-Stone, C., Sarkar, N., Antes, T., Otoc, N., Kim, Y., Jiang, Y.J. and Lu, B. (2014) A TALEN-based strategy for efficient bi-allelic miRNA ablation in human cells. *RNA*, **20**, 948-955.
18. Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S.R., Wilkie, A.O., McVean, G. and Lunter, G. (2014) Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat. Genet.*, **46**, 912-918.
19. Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Cox, A.J., Kruglyak, S. and Saunders, C.T. (2016) Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*, doi: 10.1093/bioinformatics/btv710.
20. Schurmann, N., Trabuco, L.G., Bender, C., Russell, R.B. and Grimm, D. (2013) Molecular dissection of human Argonaute proteins by DNA shuffling. *Nat. Struct. Mol. Biol.*, **20**, 818-826.
21. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.
22. Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G. *et al.* (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, **343**, 84-87.
23. Malato, Y., Naqvi, S., Schurmann, N., Ng, R., Wang, B., Zape, J., Kay, M.A., Grimm, D. and Willenbring, H. (2011) Fate tracing of mature hepatocytes in mouse liver homeostasis and regeneration. *J. Clin. Invest.*, **121**, 4850-4860.
24. Poenisch, M., Metz, P., Blankenburg, H., Ruggieri, A., Lee, J.Y., Rupp, D., Rebhan, I., Diederich, K., Kaderali, L., Domingues, F.S. *et al.* (2015) Identification of HNRNPK as regulator of hepatitis C virus particle production. *PLoS Pathog.*, **11**, e1004573.
25. Fischl, W. and Bartenschlager, R. (2013) High-throughput screening using dengue virus reporter genomes. *Methods Mol. Biol.*, **1030**, 205-219.
26. Binder, M., Quinkert, D., Bochkarova, O., Klein, R., Kezmic, N., Bartenschlager, R. and Lohmann, V. (2007) Identification of determinants involved in initiation of hepatitis C virus RNA synthesis by using intergenotypic replicase chimeras. *J. Virol.*, **81**, 5270-5283.
27. Mockenhaupt, S., Grosse, S., Rupp, D., Bartenschlager, R. and Grimm, D. (2015) Alleviation of off-target effects from vector-encoded shRNAs via codelivered RNA decoys. *Proc. Natl. Acad. Sci. USA*, **112**, E4007-4016.
28. Matsushita, T., Elliger, S., Elliger, C., Podsakoff, G., Villarreal, L., Kurtzman, G.J., Iwaki, Y. and Colosi, P. (1998) Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Ther.*, **5**, 938-945.
29. Grimm, D., Lee, J.S., Wang, L., Desai, T., Akache, B., Storm, T.A. and Kay, M.A. (2008) In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *J. Virol.*, **82**, 5887-5911.
30. Kienle, E., Senis, E., Borner, K., Niopek, D., Wiedtke, E., Grosse, S. and Grimm, D. (2012) Engineering and evolution of synthetic adeno-associated virus (AAV) gene therapy vectors via DNA family shuffling. *J. Vis. Exp.*, **2012**, 3819.
31. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G. and Tuschl, T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell*, **15**, 185-197.