

Delineation of the Exact Transcription Termination Signal for Type 3 Polymerase III

Zongliang Gao,¹ Elena Herrera-Carrillo,¹ and Ben Berkhout¹

¹Laboratory of Experimental Virology, Department of Medical Microbiology, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands

Type 3 Pol III promoters such as U6 are widely used for expression of small RNAs, including short hairpin RNA for RNAi applications and guide RNA in CRISPR genome-editing platforms. RNA polymerase III uses a T-stretch as termination signal, but the exact properties have not been thoroughly investigated. Here, we systematically measured the *in vivo* termination efficiency and the actual site of termination for different T-stretch signals in three commonly used human Pol III promoters (U6, 7SK, and H1). Both the termination efficiency and the actual termination site depend on the T-stretch signal. The T4 signal acts as minimal terminator, but full termination efficiency is reached only with a T-stretch of ≥ 6 . The termination site within the T-stretch is quite heterogeneous, and consequently small RNAs have a variable U-tail of 1–6 nucleotides. We further report that such variable U-tails can have a significant negative effect on the functionality of the crRNA effector of the CRISPR-AsCpf1 system. We next improved these crRNAs by insertion of the HDV ribozyme to avoid U-tails. This study provides detailed design guidelines for small RNA expression cassettes based on Pol III.

INTRODUCTION

In both eukaryotes and prokaryotes, small RNAs are critical for various cellular processes.¹ Based on knowledge of their mode of action, biogenesis, and processing, many small RNAs have also been developed for research purposes and therapeutic applications. For instance, the RNAi mechanism has been employed for gene regulation, which can be achieved by vector-expressed small short hairpin RNA (shRNA) or small interfering RNA (siRNA).^{2–4} The bacterial CRISPR-Cas9 system has been harnessed for genome editing in eukaryotes.^{5,6} This system requires a small guide RNA (gRNA) that guides the Cas9 nuclease for specific DNA target recognition and cleavage. Recently, the Cas9 homolog Cpf1 that uses a short CRISPR RNA (crRNA) was reported to expand the DNA-editing technology.⁷ Currently, the most popular strategy for intracellular expression of precise small RNAs is the use of RNA polymerase III (Pol III) cassettes.

Pol III transcribes short non-coding RNAs with high efficiency, including 5S rRNA (type 1), tRNAs (type 2), and other structural RNAs (type 3) such as U6 small nuclear RNA (snRNA).⁸ Type 3 genes are unique because they encode all promoter elements upstream of the transcribed region, which is ideal for the expression of RNAs of any

sequence. In addition, type 3 promoters were reported to have defined transcription start and termination sites and therefore have been widely employed for the expression of small RNA such as designed shRNAs and gRNAs. We and others have recently investigated the process of initiation of Pol III transcription at type 3 promoters,^{9,10} but the termination process has not been thoroughly studied.

Pol III termination requires a T-stretch on the non-template DNA strand without the need for other *cis*-elements or *trans*-factors.^{11,12} A recent report on yeast Pol III proposed that the non-template T-stretch is recognized by Pol III subunits, which contributes to Pol III termination.¹² Previous studies indicated that a minimum of T4 (TTTT) is required for Pol III termination in vertebrates,^{11,13} but the exact termination efficiency of human Pol III on different T-stretches was not measured. Several studies on Pol III from different species indicated that termination can occur at several sites in the T-stretch. However, the exact site of termination remains unknown because conflicting reports described a different number of U in the transcripts (≤ 3 U,¹³ 2 U,^{4,14,15} ≤ 4 U,¹⁶ 4 U,^{2,3} 4–6 U,¹⁷ ≤ 5 U,^{18,19} 5–7 U^{20,21}). Despite the massive use of type 3 Pol III promoters, detailed characterization of Pol III termination in human cells has not been performed. Here, we systematically investigate Pol III transcription termination, including the termination efficiency and the termination site using three popular human promoter systems (U6, 7SK, and H1). This study provides important information for designing Pol III-mediated small RNA expression cassettes.

RESULTS

Pol III Termination Efficiency Depends on T-Stretch Length

In order to evaluate the Pol III termination efficiency of different T-stretches, we designed U6 promoter constructs to synthesize an ~ 63 -nt transcript up to a test terminator (T1–T8) (Figures 1A and S1A). The transcripts that read through the test terminator will terminate at the T9 backup terminator at position 100. We transfected equal amounts of these constructs into HEK293T cells and isolated the total cellular RNA after 2 days. Northern blotting was performed using a probe (position 38–56) that detects both the terminated (T)

Received 22 August 2017; accepted 15 November 2017;
<https://doi.org/10.1016/j.omtn.2017.11.006>

Correspondence: Ben Berkhout, Laboratory of Experimental Virology, Department of Medical Microbiology, Academic Medical Center, University of Amsterdam, Meibergdreef 15, 1105AZ Amsterdam, the Netherlands.

E-mail: b.berkhout@amc.uva.nl



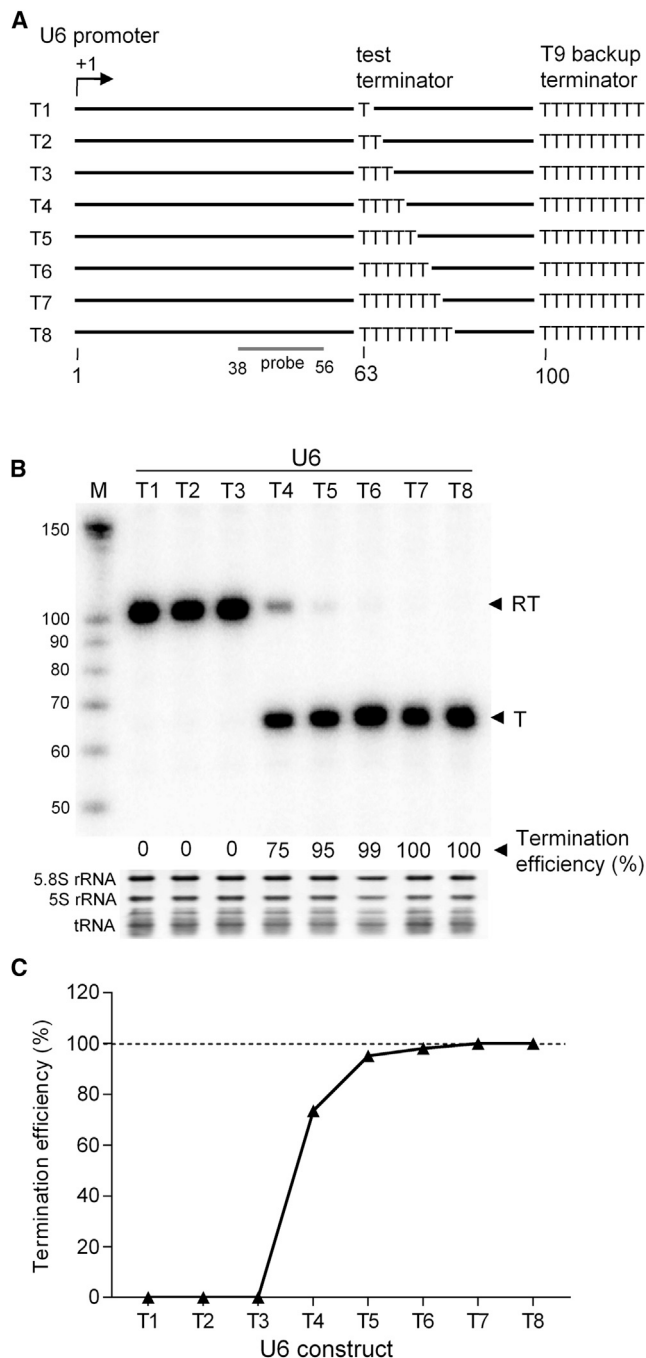


Figure 1. T-Stretch Requirements for Pol III Termination in Human U6 Promoter Cassette

(A) Schematic of U6-driven small RNA constructs. The +1 position is the Pol III transcription start site. The U6 promoter is designed to drive the ~63-nt transcripts when termination occurs at the test terminator (T1–T8). Read-through transcripts will stop at the T9 backup terminator at position 100 in all constructs due to progressive deletion of nt in the 71–77 area (position in the T1 construct). This was done to maintain the sequence context of the test terminators (see details in Figure S1A). The probe was used to detect both the short (T, terminated) and extended (RT, read-through) transcripts. (B) RNA transcript analysis by northern blotting. Equal

amounts (5 μ g) of U6 constructs were transfected into HEK293T cells, and a fixed amount (5 μ g) of isolated total cellular RNA was subjected to northern blot using the probe indicated in (A). The size (nt) of the RNA markers (lane M) is indicated. The T and RT transcripts are marked. The bands were quantitated, and the termination efficiency was determined by the following formula: Efficiency (%) = $100 \times T/(T + RT)$. The termination efficiency for different Pol III signals is shown below each lane. The rRNAs (5.8S and 5S) and tRNAs are ethidium bromide stained as loading control. (C) The termination efficiency of different Pol III signals based on data from (B). These results were reproduced in two independent experiments with similar trends.

and read-through transcripts (Figure 1B). Two major transcripts were detected, the ~100-nt read-through transcript for the T1–T3 constructs and the ~63-nt T transcript for the T4–T8 constructs, while minor ~100-nt bands were visible for T4 and to a lesser extent for T5. The RNA signals were quantitated to calculate the termination efficiency of the different terminators (Figure 1C). The T4 signals yields 75% termination efficiency and T5 reaches almost full activity (95%), while longer T-stretches result in complete termination.

The Site of Pol III Termination Is Heterogeneous

We next wanted to map the exact Pol III termination site used by the different T-stretches. For this purpose, we performed fluorescent primer extension based GeneScan analysis, which provides precise sizing, high resolution, and quantitative information on the fluorescently labeled DNA fragments.²² We selected the T4–T8 set because it represents the complete activity range from weak to full activity, and we included a control construct that is based on T1 but with substitution of the 63–100 region (see details in Figure S1). The total RNA from T4–T8 transfected HEK293T cells was isolated and ligated to a 3' adaptor, and reverse transcription was performed (Figure 2A). The FAM-labeled forward primer was used for the PCR reaction, and the resulting DNA products, together with a size marker, were subjected to GeneScan analysis. The total RNA from the control transfection was also subjected to the GeneScan procedure without the 3' adaptor ligation step. This produced a single peak signal of the expected size, which corresponds to the DNA derived from Pol III transcripts that terminate immediately at the first T (T¹) of the T-stretch (Figure 2B; see Figure S1 for details). Densitograms obtained for the different termination sites were plotted (Figure 2B). Multiple signals that are a few nt longer than the control signal were observed for the T4–T8 constructs, but termination always occurred within the T-stretch. These results show that Pol III termination occurs at multiple sites within the T-stretch. Signals were quantitated, and the percentage of termination at each T position was calculated (Figure 2C). A broad distribution of termination peaks was observed. Termination at T¹ was observed, but at a minimal efficiency, which reaches a maximum of 5% for the T4 construct that has only three downstream alternative termination sites. T4 terminates mostly at T^{3–4}, and T5 stops largely at T^{3–5}. The T6–T8 constructs exhibit a similar termination profile and also prominently terminate at T^{3–5}, with some low-level termination at the T² and T⁶ positions. In summary, Pol III terminates at variable sites within a T-stretch of ≥ 4 , but most events occur in the T³–T⁵ window.

amounts (5 μ g) of U6 constructs were transfected into HEK293T cells, and a fixed amount (5 μ g) of isolated total cellular RNA was subjected to northern blot using the probe indicated in (A). The size (nt) of the RNA markers (lane M) is indicated. The T and RT transcripts are marked. The bands were quantitated, and the termination efficiency was determined by the following formula: Efficiency (%) = $100 \times T/(T + RT)$. The termination efficiency for different Pol III signals is shown below each lane. The rRNAs (5.8S and 5S) and tRNAs are ethidium bromide stained as loading control. (C) The termination efficiency of different Pol III signals based on data from (B). These results were reproduced in two independent experiments with similar trends.

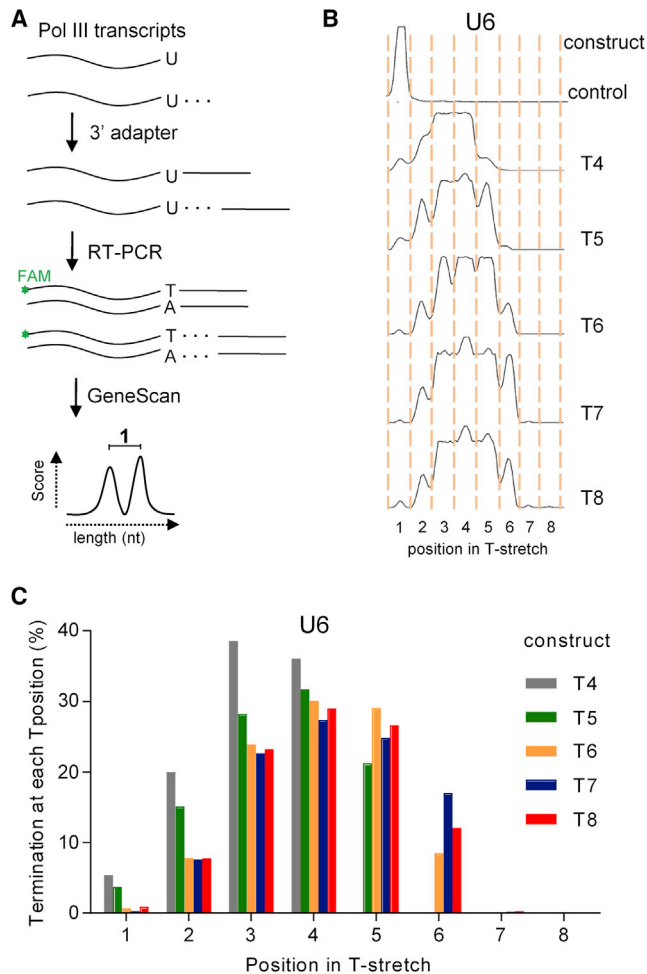


Figure 2. Mapping the Pol III Termination Site in the U6 System

(A) The GeneScan procedure to map the termination site. Two days after transfection of HEK293T cells, total RNA was harvested. Total cellular RNA was ligated to a 3' adaptor and then reverse transcribed using an adaptor primer. The 5' FAM forward primer and the adaptor primer are subsequently used for PCR, and the resulting products together with a size marker are subjected to GeneScan, which measures the FAM fluorescence and provides a signal with a resolution of 1 nt. (B) The GeneScan output. The control yields a single peak of the size expected for the termination at the T1 position. Each peak represents the termination activity at the corresponding T position. (C) Quantitation of the signals in (B). All results were reproduced in two independent experiments with similar trends.

Similar Pol III Termination Profiles in 7SK and H1 Cassettes

We next wanted to test the termination profile of other type 3 Pol III complexes. To do so, the U6 promoter was replaced by the 7SK and H1 promoters. Northern blot analysis showed read-through transcripts for the T1–T4 constructs and terminated transcripts for T5–T8 (Figures 3A and 3B). The termination efficiency was calculated and plotted together with that of the U6 constructs to allow a direct comparison (Figure 3C). Similar to the results with U6, the 7SK and H1 constructs do not induce termination at T1–T3 stretches and reach full termination efficiency at the T6–T8 stretches. Some variation between the

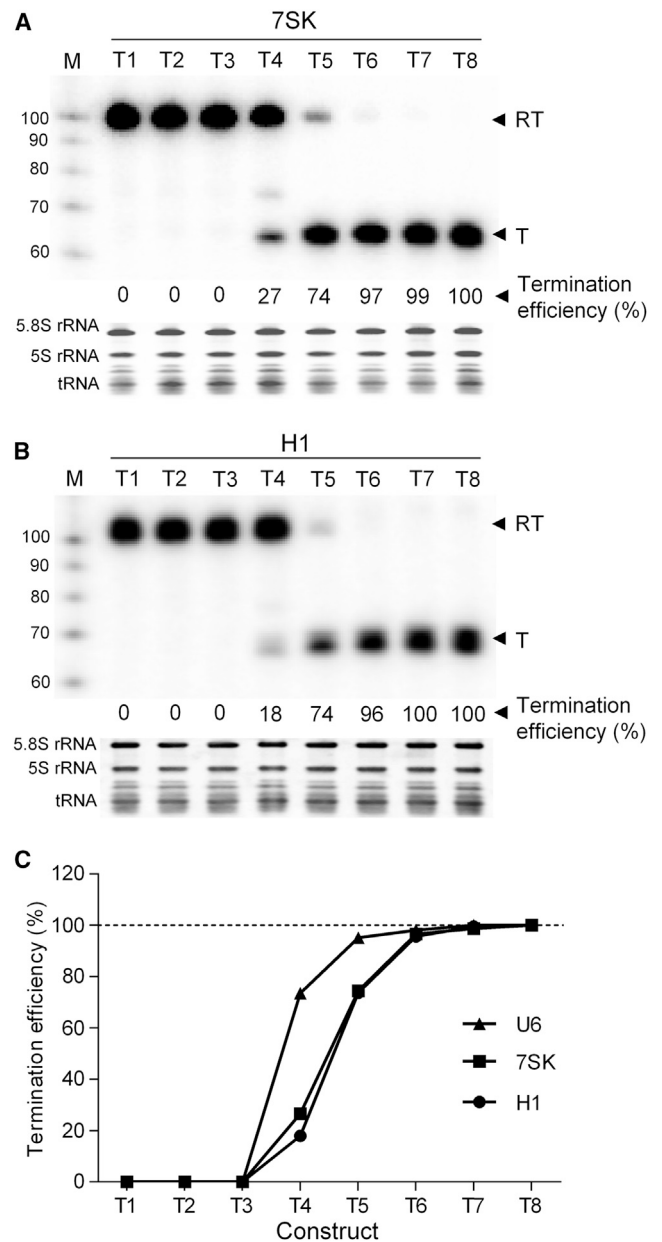


Figure 3. T-Stretch Termination Efficiency in the 7SK and H1 Systems

The 7SK and H1 promoters were designed to replace the promoter in the U6 constructs. See Figure 1A for further details. Northern blotting was performed to evaluate the termination efficiency of the 7SK (A) and H1 (B) systems. The calculated termination efficiency for different Pol III signals is shown below each lane. (C) The termination efficiency profile of 7SK and H1 was calculated and plotted together with the U6 data from Figure 1C. All data were reproduced in two independent experiments with similar trends.

different Pol III complexes was observed for the intermediately active T4 and T5 signals. T4 is only 27% and 18% active in the 7SK and H1 context but reaches 75% efficiency for U6. T5 reaches 74% activity for 7SK and H1 but is almost fully active in the U6 context.

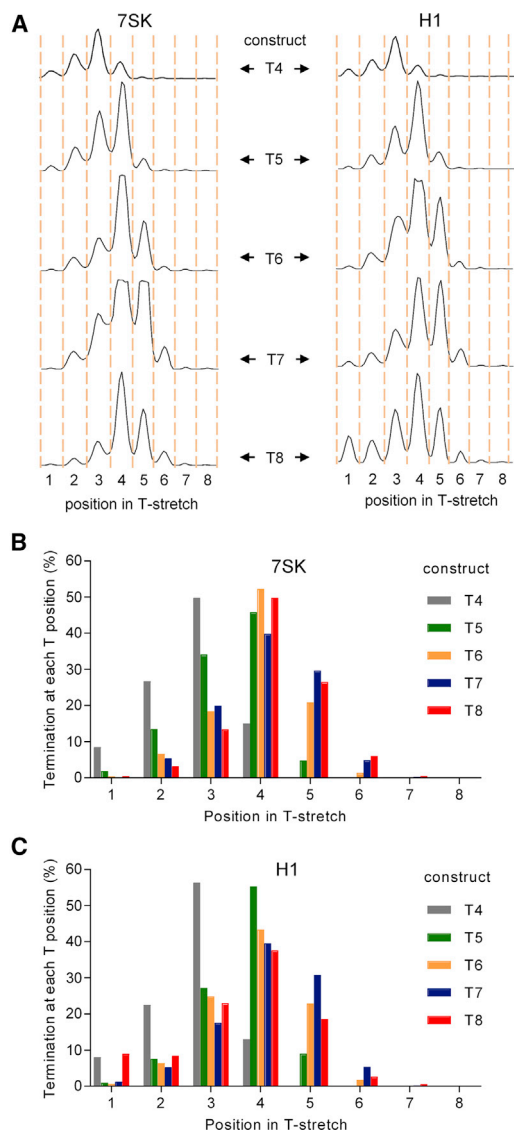


Figure 4. Mapping the Pol III Termination Site in the 7SK and H1 Systems (A) GeneScan output representing the Pol III termination profile (7SK, left; H1, right). The GeneScan was performed and the termination site profiles of T4–T8 constructs were plotted as in Figure 2. Each peak represents the termination activity at the corresponding T position. (B and C) The signals in (A) were quantitated, and the termination efficiency at each T position was calculated (7SK, B; H1, C). All reported data were reproduced in two independent experiments with similar trends.

The termination sites were subsequently mapped by GeneScan, and a similar profile was obtained for 7SK and H1 (Figure 4A; quantification in Figures 4B and 4C). T4 and T5 terminate mostly at T²⁻³ and T³⁻⁴, respectively, with minor activity in the flanking nt. T6–T8 terminates largely at T³⁻⁵ with a prominent peak at T⁴. Thus, the termination site profiles of 7SK and H1 differ slightly from that of the U6 cassette, but termination heterogeneity is a common property.

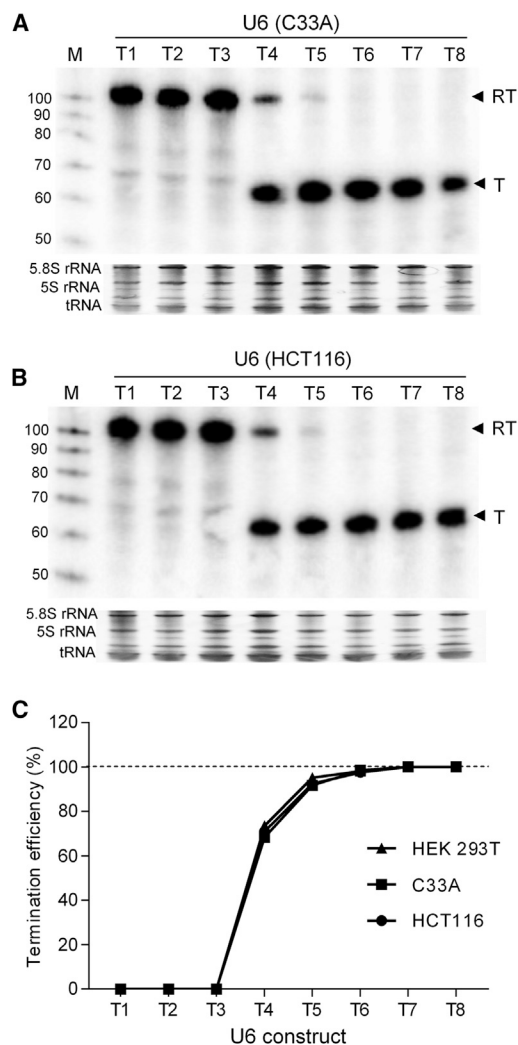


Figure 5. Pol III Termination Is Cell Type Independent

(A and B) The U6 constructs (T1–T8) were transfected into C33A cells (A) and HCT116 cells (B), and northern blot analysis was performed as described above. (C) The termination efficiency in these two cell lines was calculated as described in Figure 1 and plotted together with the HEK293T data in Figure 1C. All results were reproduced in two independent experiments with similar trends.

The Pol III Termination Profiles Are Cell Type Independent

To test if these distinct Pol III termination properties differ across cell types, two additional cell lines were selected for a test of the U6 constructs. Total cellular RNA was isolated 2 days after transfection, and Northern blot analysis was performed. Extremely similar patterns were observed for C33A cells (Figure 5A) and HCT116 cells (Figure 5B) when compared to the original HEK293T cells (Figure 1B). The termination efficiency profiles in all three cell types exhibited identical trends (Figure 5C), indicating that Pol III termination is independent of co-factors that vary in concentration or activity among these cell types.

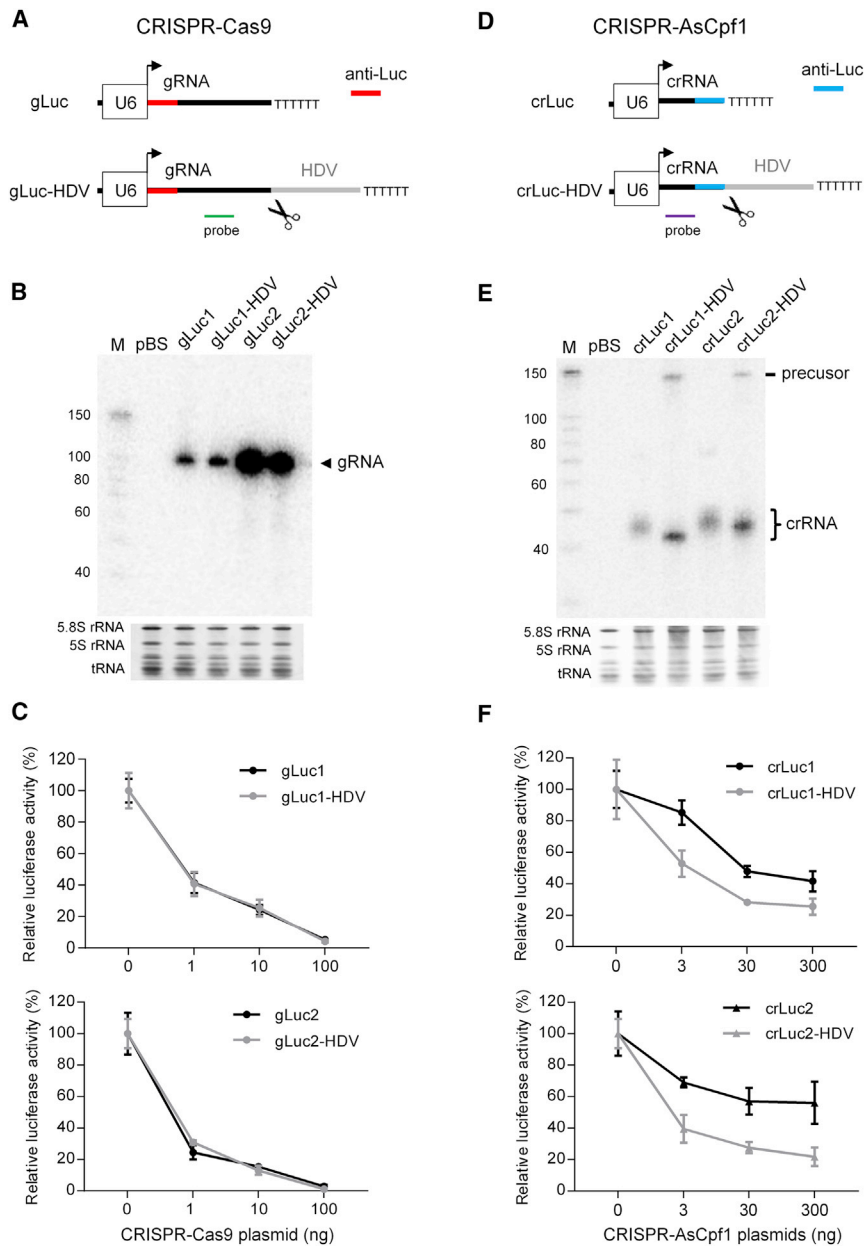


Figure 6. The Effect of Variable U-Tail on gRNA and crRNA Activity in the CRISPR-Cas9 and CRISPR-AsCpf1 Systems

(A) Schematic of the gRNA expression cassettes. Construct gLuc uses the U6 promoter to transcribe the anti-Luc gRNA followed by the T6 terminator. The gLuc-HDV cassette encodes a HDV ribozyme between the gRNA and the T6 signal (see Figure S2 for details). The HDV ribozyme was designed to induce self-cleavage at the gRNA-HDV border (marked as scissor). The anti-Luc sequence of gLuc is marked in red; the scaffold is in black. The probe targeting the gRNA scaffold was used to detect gRNA expression. (B) An equal amount of the gLuc or gLuc-HDV constructs was transfected into HEK293T cells, and a fixed amount of total cellular RNA was used for northern blot analysis. The pBluescript SK (pBS) plasmid was used as negative control. M is the RNA marker for size estimation. The position of gRNA was indicated as arrowhead. (C) The anti-Luc CRISPR-Cas9 plasmids were titrated (1, 10, or 100 ng) and co-transfected with 100 ng Luc reporter and 2 ng Renilla plasmid. Luc activity was determined 2 days after transfection. The Luc activity without CRISPR-Cas9 plasmid was set as 100%. The results are shown as the mean value \pm SD ($n = 3$). (D) Schematic of the crRNA expression cassettes. See (A) for further details. The anti-Luc sequence of crLuc is marked in blue. (E) Northern blot analysis with the probe indicated in (D). (F) The CRISPR-Cpf1 plasmids were titrated (3, 30, or 300 ng) for Luc knockdown assay. The mass ratio of crRNA- and Cpf1-expressing plasmids was kept at 1:2. All data are presented as the mean value \pm SD ($n = 3$).

between the gRNA and T6 signal to create the novel gLuc-HDV construct. The HDV ribozyme forms a specific tertiary RNA conformation that triggers self-cleavage immediately at the gRNA border (scissor in Figure 6A). Two anti-Luc gRNAs were designed and tested (see Table 1).

To evaluate the gRNA expression level, these four anti-Luc CRISPR-Cas9 constructs were transfected into HEK293T cells, and the total cellular RNA was isolated and subjected to northern blotting using a probe targeting the gRNA scaffold sequence (Figure 6A). All four constructs pro-

duce a gRNA transcript of similar size around 100 nt that is the predicted gRNA size, demonstrating effective cleavage by the HDV ribozyme (Figure 6B). The two gLuc2 constructs produce more transcript than the two gLuc1 constructs, but the addition of the HDV ribozyme did not influence the RNA production level. This difference may be caused by sequence differences around the transcription initiation area that acts as a key determinant for transcription efficiency.^{9,10} Alternatively, the gLuc2 transcript may be more stable than gLuc1.

To assess the DNA cleavage efficiency of these CRISPR-Cas9 constructs, they were titrated in a co-transfection with the Luc reporter,

The Effect of a Variable 3' End U-Tail on gRNA and crRNA Activity

A T-stretch of minimally six residues causes efficient termination of Pol III transcription and thus should be used for optimal expression in small RNA cassettes. However, a variable number of U residues will be transcribed into the 3' end of these RNAs (Figures 2 and 4), which may have an effect on their function and activity. We first tested this potential effect for a gRNA of the CRISPR-Cas9 system. Two gRNA expression constructs were made (Figure 6A). The standard gLuc construct uses the U6 promoter to transcribe an anti-Luc gRNA followed by an efficient T6 terminator. To eliminate the U-tail of variable length, we inserted the hepatitis delta virus (HDV) ribozyme^{23,24}

Table 1. Luc Target Sequence (PAM Underlined)

Name	Target Sequence (5'-3') and PAM
gLuc1	AACGGCGGCGGAAGTTCACCGG
gLuc2	GGAGCACGGAAAGACGATGACCGG
crLuc1	TTTATAATGAACGTGAATTGCTCAACA
crLuc2	TTTCAGTCGATGTACACGTTCGTCACA

and a Renilla plasmid was included to control for the transfection efficiency. The relative Luc activity (Luc/Renilla) was calculated, and the Luc activity obtained in the absence of CRISPR-Cas9 construct was set at 100%. All anti-Luc gRNAs mediated potent Luc inhibition (Figure 6C). gLuc2 was more active than gLuc1, which may relate to higher gLuc2 expression. Most importantly, similar knockdown was scored for gLuc with or without the HDV ribozyme. Thus, the presence of a variable 3' terminal U-tail does not seem to have an effect on the functionality of these gRNAs in the CRISPR-Cas9 system.

We next used the same approach to test the effect of a variable 3' U-tail on crRNA molecules of the CRISPR-AsCpf1 system. crLuc and crLuc-HDV constructs were made (Figure 6D), again with two anti-Luc sequences (see Table 1). The crRNA expression was detected by northern blotting with the probe targeting the As-crRNA scaffold. Transcripts corresponding to the crRNA of ~43 nt were detected for all four constructs, indicating efficient cleavage of the HDV ribozyme (Figure 6E). Minor bands reflecting the crRNA-HDV precursor were detected for the two crRNA-HDV constructs. Such precursors were not observed for the gRNA-HDV construct. Perhaps HDV misfolding is more prominent in the crRNA context that is less structured than the gRNA transcript. The two crLuc transcripts are a bit longer and more diffuse than the corresponding crLuc-HDV transcripts, consistent with the presence of a U-tail of variable length. Unlike the gRNA results in Figure 6B, the crRNA expression level is roughly similar for the four constructs (Figure 6E). But unlike the gRNA constructs, all crRNA constructs have the same crRNA scaffold sequence proximal to the transcription initiation area, which is a key determinant of the transcription efficiency.^{9,10} Luc knockdown was performed with increasing amounts of the CRISPR-AsCpf1 plasmids. A noticeable difference between the crLuc and crLuc-HDV constructs was apparent for both two anti-Luc agents (Figure 6F). The more potent Luc inhibition by crLuc-HDV than crRNA indicates that the U-tail can have a significant negative effect on the crRNA activity.

DISCUSSION

In this study, we systematically investigated details of Pol III transcription termination in human cells using three popular human promoter systems (U6, 7SK, and H1) that are widely used for small RNA synthesis.^{3,7,25} The termination efficiency and the actual site of termination vary depending on the T-stretch signal. The minimal T4 terminator caused incomplete termination, and complete termination requires a signal of $\geq 6T$. The site of termination is heterogeneous and thus a variable number of U residues will be

generated at the 3' end of the transcribed RNA, which is a common property of the three Pol III-cassettes tested. These new insights provide useful guidelines for designing optimal Pol III-driven small RNA units.

Previous studies showed that the Pol III termination efficiency by the short T4 signal is influenced by the flanking sequence, but this effect was reduced for the more active T5 termination signal.^{11,26,27} Thus, the termination efficiency determined by us for the shorter T4 and T5 signals may vary with different flanking sequences. A T5-T7 stretch is the most commonly used terminator for Pol III-driven small RNA expression units.^{17,25,28} However, we show that T5 induces incomplete Pol III termination in a cell type-independent manner and therefore reduces the level of small RNA expression. In addition, the failure of complete termination may produce a low level of 3' extended RNAs with unwanted activity. The ongoing Pol III transcription may also interfere with the expression of downstream genes. Thus, a minimal T6 signal that achieves full Pol III termination should be used in small RNA expression cassettes.

Due to the heterogeneous termination site, a variable U-tail is present at the 3' end of Pol III-transcribed RNA. The effect of this U-tail was tested on the gRNA and crRNA activity of two CRISPR gene-editing systems. No effect was scored for two gRNAs, but a profound negative effect was measured for two crRNA molecules that target the same Luc target. This differential sensitivity of gRNA and crRNA molecules to the variable U-tail may be linked to differences in their structure. In gRNAs, the scaffold sequence required for Cas9 binding is located adjacent to the U-tail, while for crRNAs the protospacer sequence that is involved in target DNA sequence recognition is located near the U-tail. It therefore seems that the U-tail with variable length does not affect Cas9 binding to the gRNA but does affect target recognition and activity of the crRNA. Our results indicate that a variable U-tail can affect the functionality of some small RNAs, and caution should be taken when Pol III cassettes are used for synthesizing small RNAs. For instance, the widely used shRNA design has a 3' UU overhang based on the miRNA-processing pathway.^{29,30} However, our study indicates that the frequently used Pol III systems generate variable U-tails, which may affect the activity. Similar effect can be expected for the constructs that express miRNAs or modified shRNAs like AgoshRNA molecules.^{30,31}

The self-cleaving HDV ribozyme has previously been used for the expression of the gRNAs.²⁴ Most importantly, such RNA processing element can be used as a strategy to generate multiple gRNAs from a single transcript to allow multiplex gene regulation.³² Here, we demonstrated that HDV insertion does not improve gRNA activity, but the crRNA activity was significantly enhanced. Such improvement was scored for two independent crRNAs that target the Luc DNA. We therefore propose that the crRNA-HDV context in Pol III cassette should ideally be used to increase the activity of the CRISPR-AsCpf1 system. This optimization is very welcome as the CRISPR-AsCpf1 system is only moderately active in our hands (Z. Gao, unpublished data).

Table 2. Probes for Northern Blot

Name	Target Sequence (5'-3') and PAM
Pol47	GTGAAGGGGCAGTAGTAAT
gRNA	CAAGTTGATAACGGACTAG
crRNA	TCTACAAGAGTAGAAATTA

Thus, the HDV ribozyme can be used to eliminate the potential negative effect of U-tail variation caused by Pol III termination or multiplex small RNAs expression, but caution should be taken because the cleavage efficiency of HDV may vary in different sequence contexts.

The three Pol III promoter cassettes tested in this study generally show similar termination profiles. However, one surprising difference is that termination at the sub-optimal T4 and T5 signals in the U6 system is much more efficient than that in the 7SK and H1 systems. All tested systems produce complete termination when the T-stretch is ≥ 6 . The molecular mechanism behind this difference remains unknown, but the U6-recruited Pol III transcription complex seems more prone to termination. A sequence comparison indicated that all three promoters have common motifs for recruitment of Pol III, but with significant sequence variation.^{33–35} Such sequence differences may contribute to variation in transcription factor recruitment and affect the sensitivity of the Pol III transcription complex toward termination.

MATERIALS AND METHODS

Plasmid Construction

The vectors pSilencer 2.0-U6 (Ambion), psiRNA-h7SK hygro G1 (Invivogen), and pSUPER (OligoEngine) were used as source for the three human Pol III promoters (U6, 7SK, and H1, respectively). To generate T1–T8 constructs (Figures 1A and S1A), the DNA oligo nucleotides with different T-stretch signals were annealed and inserted into three vectors using the proper restriction enzyme sites (BamHI, HindIII for U6, Acc65I, HindIII for 7SK and BglII, HindIII for H1). The CRISPR vectors pX458 (#48138, Addgene) and pcDNA3.1-hAsCpf1 (#69982, Addgene) were kindly donated by Feng Zhang.^{7,25} The pX458 plasmid contains a human U6 promoter-mediated gRNA expression cassette and a “human codon-optimized” *Streptococcus pyogenes* Cas9 expression cassette. The pcDNA3.1-hAsCpf1 vector expresses the “human codon-optimized” AsCpf1 nuclease (from *Acidaminococcus* sp). The pSilencer 2.0-U6 vector with the human U6 promoter was used for crRNA expression. The DNA oligonucleotides encoding anti-Luc sequences in gLuc and crLuc were annealed and inserted into the pX458 (BbsI sites) and pSilencer 2.0-U6 (BamHI and HindIII sites) vectors, respectively. The U6-gLuc-HDV and U6-crRNA-HDV gene fragments were synthesized by Integrated DNA technology (IDT) and cloned into pX458 (AflIII and XbaI sites) and pSilencer 2.0-U6 (PmlI and HindIII sites) by Gibson cloning according to the manufactures’ instructions (New England Biolabs). All vectors were verified by sequencing using the BigDye Terminator v1.1 Cycle Sequencing Kit (ABI).

Cell Culture

HEK293T cells, C33A cells, and HCT116 cells were cultured in DMEM (Life Technologies, Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal calf serum (FCS), penicillin (100 U/mL), and streptomycin (100 μ g/mL). C33A is a human cervical cancer cell line, and HCT116 is a human colon cancer cell line. Cells were trypsinized and seeded 1 day prior to transfection.

Dual-Luciferase Reporter Assay

CRISPR constructs were co-transfected (titration as indicated in Figures 6C and 6F) into HEK293T cells with 100 ng of pGL3 Luc reporter and 2 ng Renilla plasmid using lipofectamine 2000 (Invitrogen) according to the manufactures’ instructions. Two days post-transfection, luciferase activity was measured with the Dual-Luciferase Reporter Assay System (Promega, Madison, WI, USA) according to the manufacturer’s protocol. The relative Luc activity was determined by the Luc/Renilla ratio. The results were corrected for between-session variation as described previously.³⁶

Northern Blotting

Northern blotting was performed as previously described.¹⁰ In brief, 1.5×10^6 HEK293T cells per 25 cm³ flask were transfected with equimolar quantities (5 μ g) of the constructs using lipofectamine 2000 (Invitrogen). Total cellular RNA was harvested 2 days post-transfection using the mirVana miRNA isolation kit (Ambion). Equal amount (5 μ g) of total RNA was electrophoresed in a 15% denaturing polyacrylamide gel (Precast Novex TBU gel, Life Technologies). [γ -³²P]-labeled decade RNA marker (Life Technologies) was run alongside for size estimation. To check for equal sample loading, the gel was stained in 2 μ g/mL ethidium bromide and visualized under UV light. RNA in the gel was electro-transferred to a positively charged nylon membrane (Boehringer Mannheim, GmbH). The locked nucleic acid (LNA) oligonucleotide probes (see Table 2) were 5’ end-labeled with [γ -³²P]-ATP (0.37 MBq/ μ L, Perkin Elmer) using the kinaseMax kit (Ambion). The blots were incubated with the labeled probe in 10 mL ULTRAhyb hybridization buffer (Ambion) at 42°C for overnight. The membranes were washed twice for 5 min at 42°C with $2 \times$ SSC/0.1% SDS and twice for 5 min at 42°C with $0.1 \times$ SSC/0.1% SDS. The signals were captured by the Typhoon FLA 9500 (GE Healthcare Life Sciences) and analyzed by ImageJ software.

Mapping Pol III Termination Sites by Fluorescent Primer

Extension

Two hundred nanograms of total cellular RNA from transfected cells was ligated to an adenylated 3’-adaptor (5’-(rApp)-GGAACCAT CAATATCTCGTATGCCGTCTTCTGCTTG-(3ddC)-3’) (IDT) by a truncated T4 RNA Ligase 2 (New England Biolabs). The RNA-adaptor product was reverse-transcribed using ThermoScript RT-PCR System (Invitrogen) with the specific adaptor primer 5’-CAAG CAGAAGACGGCATAACG-3’. PCR amplification was performed with Phusion High-Fidelity DNA Polymerases (New England Biolabs) that produces PCR products with blunt ends (without A addition). The PCR reaction was run using two primers: 5’-(FAM)-GA

TATCACCGGTATATTAAC-3' and 5'-CAAGCAGAAGACGGCA TACG-3' according to the manufacturer's instructions. The total RNA from the control constructs was used for RT-PCR without the ligation step (Figure S1B). 1 μ L of the PCR products together with 1.5 μ L Rox 500 Size Standard were run on the ABI PRISM 3010 XL Genetic Analyzer (Applied Biosystems) with default parameters. The output data were analyzed using GeneMapper software v4.0 (Applied Biosystems), and the termination profiles were calibrated against the size of the control sample.

SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures and can be found with this article online at <https://doi.org/10.1016/j.omtn.2017.11.006>.

AUTHOR CONTRIBUTIONS

Z.G., E.H.-C., and B.B. designed the experiments. Z.G. and B.B. drafted the manuscript. Z.G. conducted the experiments, and all authors analyzed the data.

CONFLICTS OF INTEREST

No conflicts of interest were disclosed.

ACKNOWLEDGMENTS

This research was supported by NWO-Chemical Sciences (TOP grant) and ZonMW (Translational Gene Therapy program). Z.G. is supported by a scholarship from the China Scholarship Council (CSC). We thank Yi Zheng for the kind donation of reagents.

REFERENCES

- Morris, K.V., and Mattick, J.S. (2014). The rise of regulatory RNA. *Nat. Rev. Genet.* *15*, 423–437.
- Yu, J.-Y., DeRuiter, S.L., and Turner, D.L. (2002). RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc. Natl. Acad. Sci. USA* *99*, 6047–6052.
- Miyagishi, M., and Taira, K. (2002). U6 promoter-driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat. Biotechnol.* *20*, 497–500.
- Brummelkamp, T.R., Bernards, R., and Agami, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells. *Science* *296*, 550–553.
- 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.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816–821.
- Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* *163*, 759–771.
- Schramm, L., and Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. *Genes Dev.* *16*, 2593–2620.
- Ma, H., Wu, Y., Dang, Y., Choi, J.-G., Zhang, J., and Wu, H. (2014). Pol III promoters to express small RNAs: delineation of transcription initiation. *Mol. Ther. Nucleic Acids* *3*, e161.
- Gao, Z., Harwig, A., Berkhout, B., and Herrera-Carrillo, E. (2017). Mutation of nucleotides around the +1 position of type 3 polymerase III promoters: the effect on transcriptional activity and start site usage. *Transcription*, e1322170.
- Arimbasseri, A.G., Rijal, K., and Maraia, R.J. (2013). Transcription termination by the eukaryotic RNA polymerase III. *Biochim. Biophys. Acta* *1829*, 318–330.
- Arimbasseri, A.G., and Maraia, R.J. (2015). Mechanism of transcription termination by RNA polymerase III utilizes a non-template strand sequence-specific signal element. *Mol. Cell* *58*, 1124–1132.
- Maraia, R.J., and Lamichhane, T.N. (2011). 3' processing of eukaryotic precursor tRNAs. *Wiley Interdiscip. Rev. RNA* *2*, 362–375.
- Lee, N.S., Dohjima, T., Bauer, G., Li, H., Li, M.-J., Ehsani, A., Salvaterra, P., and Rossi, J. (2002). Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat. Biotechnol.* *20*, 500–505.
- Bogenhagen, D.F., and Brown, D.D. (1981). Nucleotide sequences in *Xenopus* 5S DNA required for transcription termination. *Cell* *24*, 261–270.
- Paul, C.P., Good, P.D., Winer, I., and Engelke, D.R. (2002). Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* *20*, 505–508.
- Boden, D., Pusch, O., Lee, F., Tucker, L., Shank, P.R., and Ramratnam, B. (2003). Promoter choice affects the potency of HIV-1 specific RNA interference. *Nucleic Acids Res.* *31*, 5033–5038.
- Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W.S., Karpilow, J., and Khvorova, A. (2005). The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA* *11*, 674–682.
- Sui, G., Soohoo, C., Affar, B., Gay, F., Shi, Y., Forrester, W.C., and Shi, Y. (2002). A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl. Acad. Sci. USA* *99*, 5515–5520.
- Matsuzaki, H., Kassavetis, G.A., and Geiduschek, E.P. (1994). Analysis of RNA chain elongation and termination by *Saccharomyces cerevisiae* RNA polymerase III. *J. Mol. Biol.* *235*, 1173–1192.
- Shaaban, S.A., Bobkova, E.V., Chudzki, D.M., and Hall, B.D. (1996). In vitro analysis of elongation and termination by mutant RNA polymerases with altered termination behavior. *Mol. Cell. Biol.* *16*, 6468–6476.
- Fekete, R.A., Miller, M.J., and Chatteraj, D.K. (2003). Fluorescently labeled oligonucleotide extension: a rapid and quantitative protocol for primer extension. *Biotechniques* *35*, 90–94, 97–98.
- Nakano, S., Chadalavada, D.M., and Bevilacqua, P.C. (2000). General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science* *287*, 1493–1497.
- Gao, Y., and Zhao, Y. (2014). Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J. Integr. Plant Biol.* *56*, 343–349.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* *8*, 2281–2308.
- Goodier, J.L., and Maraia, R.J. (1998). Terminator-specific recycling of a B1-Alu transcription complex by RNA polymerase III is mediated by the RNA terminus-binding protein La. *J. Biol. Chem.* *273*, 26110–26116.
- Braglia, P., Percudani, R., and Dieci, G. (2005). Sequence context effects on oligo(dT) termination signal recognition by *Saccharomyces cerevisiae* RNA polymerase III. *J. Biol. Chem.* *280*, 19551–19562.
- Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X., and Skarnes, W.C. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat. Methods* *11*, 399–402.
- Bofill-De Ros, X., and Gu, S. (2016). Guidelines for the optimal design of miRNA-based shRNAs. *Methods* *103*, 157–166.
- Lam, J.K., Chow, M.Y., Zhang, Y., and Leung, S.W. (2015). siRNA versus miRNA as therapeutics for gene silencing. *Mol. Ther. Nucleic Acids* *4*, e252.
- Herrera-Carrillo, E., Gao, Z.L., Harwig, A., Heemskerck, M.T., and Berkhout, B. (2017). The influence of the 5'-terminal nucleotide on Ago2-RNA activity and

- biogenesis: importance of the polymerase III transcription initiation site. *Nucleic Acids Res.* 45, 4036–4050.
32. Xu, L., Zhao, L., Gao, Y., Xu, J., and Han, R. (2017). Empower multiplex cell and tissue-specific CRISPR-mediated gene manipulation with self-cleaving ribozymes and tRNA. *Nucleic Acids Res.* 45, e28.
 33. James Faresse, N., Canella, D., Praz, V., Michaud, J., Romascano, D., and Hernandez, N. (2012). Genomic study of RNA polymerase II and III SNAPc-bound promoters reveals a gene transcribed by both enzymes and a broad use of common activators. *PLoS Genet.* 8, e1003028.
 34. Murphy, S., Tripodi, M., and Melli, M. (1986). A sequence upstream from the coding region is required for the transcription of the 7SK RNA genes. *Nucleic Acids Res.* 14, 9243–9260.
 35. Dykxhoorn, D.M., Novina, C.D., and Sharp, P.A. (2003). Killing the messenger: short RNAs that silence gene expression. *Nat. Rev. Mol. Cell Biol.* 4, 457–467.
 36. Ruijter, J.M., Thygesen, H.H., Schoneveld, O.J.L.M., Das, A.T., Berkhout, B., and Lamers, W.H. (2006). Factor correction as a tool to eliminate between-session variation in replicate experiments: application to molecular biology and retrovirology. *Retrovirology* 3, 2.