Citation: Molecular Therapy—Nucleic Acids (2013) 2, e104; doi:10.1038/mtna.2013.33

www.nature.com/mtna

An Algorithm for Generating Small RNAs Capable of Epigenetically Modulating Transcriptional Gene Silencing and Activation in Human Cells

Amanda Ackley[1], Alexandra Lenox[1], Kenneth Stapleton[1], Stuart Knowling[1], Tim Lu[1], Kenny S Sabir[2], Peter K Vogt[1] and Kevin V Morris[1,3]

Small noncoding antisense RNAs (sasRNAs) guide epigenetic silencing complexes to target loci in human cells and modulate gene transcription. When these targeted loci are situated within a promoter, long-term, stable epigenetic silencing of transcription can occur. Recent studies suggest that there exists an endogenous form of such epigenetic regulation in human cells involving long noncoding RNAs. In this article, we present and validate an algorithm for the generation of highly effective sasRNAs that can mimic the endogenous noncoding RNAs involved in the epigenetic regulation of gene expression. We validate this algorithm by targeting several oncogenes including *AKT-1*, *c-MYC*, *K-RAS*, and *H-RAS*. We also target a long antisense RNA that mediates the epigenetic repression of the tumor suppressor gene *DUSP6*, silenced in pancreatic cancer. An algorithm that can efficiently design small noncoding RNAs for the epigenetic transcriptional silencing or activation of specific genes has potential therapeutic and experimental applications.

Molecular Therapy—Nucleic Acids (2013) **2**, e104; doi:10.1038/mtna.2013.33; published online 9 July 2013
Subject Category: siRNAs, shRNAs, and miRNAs

Introduction

Recent studies have revealed a mechanism of transcriptional regulation involving long, noncoding RNAs. These RNAs are antisense to protein-coding genes and affect transcription by epigenetic modification of the chromatin.[1] Parts of this endogenous regulatory pathway overlap with steps known from transcriptional gene silencing (TGS) with small noncoding antisense RNAs (sasRNAs) or small interfering RNAs (siRNAs).[2] SasRNAs, designed to target promoter sequences, can induce silent-state epigenetic changes at the targeted loci, mediating long-term, stable TGS.[3] These small RNAs are exogenous to the cell, but there is also emerging evidence that their mechanisms of action involve endogenous RNA-mediated transcriptional controls.[4–6] Intervention in such endogenous mechanisms could have important medical applications.

Transcriptional controls are important in cancer. The gain of function in oncogenes can result from transcriptional upregulation, and loss of function in tumor suppressor genes is often due to transcriptional silencing.[7–9] Silenced genes typically show characteristic epigenetic marks at their promoters.[10,11] Some tumor suppressor genes, such as *p15*, *p21*, and *PTEN*, are regulated by endogenous antisense transcripts.[4–6] The epigenetic silencing of *p21* and *PTEN*[4,6] is mechanistically similar to TGS induced by promoter-targeting sasRNAs.[2] These observations suggest the existence of endogenous RNA-mediated transcriptional regulation that works through epigenetic modification and involves natural antisense transcripts.[1,12,13]

To date, however, while there are extensive siRNA design algorithms for RNA interference-mediated gene silencing,[14,15] the design of promoter-targeting sasRNAs, which would mimic the endogenous natural antisense transcript pathway, has not been developed. We surmised, based on previous observations,[16,17] that TGS works through an RNA–RNA interaction at the targeted promoter. It may then be possible to rationally design sasRNAs capable of inducing TGS in human cells.[16,18] In this article, we present an algorithm for the identification of small RNAs that can mediate TGS. We include examples of TGS of several oncogenes and TGS that affects the promoter of an endogenous natural antisense transcript. Effective TGS operating on the promoter of a natural antisense transcript can elevate the transcription from the sense promoter, offering a method to activate gene expression.

Results

Small RNAs effective in the induction of TGS

SasRNAs, siRNAs, or short hairpin RNAs can induce TGS of various gene promoters in human cells.[2] However, not all small RNAs show this effect on transcription. We compared effective versus noneffective small RNAs from published literature in order to identify shared features that are correlated with transcriptional regulator activity (**Supplementary Tables S1** and **S2**, respectively). Small RNAs that function in initiating TGS typically show a string of purines from ~4 to 12 bp in the region targeting the promoter sequence (**Supplementary Table S1**). Long noncoding RNAs, which are the endogenous drivers of TGS in human cells, also

[1]Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, California, USA; [2]Garvan Institute of Medical Research, Sydney, Australia; [3]School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, Australia. Correspondence: Kevin V Morris, Department of Molecular and Experimental Medicine, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037, USA. E-mail: kmorris@scripps.edu or Peter K Vogt, Department of Molecular and Experimental Medicine, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037, USA. E-mail: pkvogt@scripps.edu
Keywords: epigenetic; non-coding RNA; oncogene; transcription; tumor suppressor gene
Received 15 May 2013; accepted 15 May 2013; advance online publication 9 July 2013. doi:10.1038/mtna.2013.33

contain strings of purines, suggesting it is of biological relevance in RNA-directed control of gene expression.[19,20] We speculated that this purine string was mechanistically relevant for TGS. In order to examine this possibility, we produced a computer algorithm taking into account the observed strings of purines as well as GC pairing (**Figure 1a**). In using this algorithm (http://www.morrislab.unsw.edu.au/TGSalgorithm/Algorithm/purine_finder.html), sequences from a desired target locus, generally a gene promoter, are copied and pasted into the "paste sequence" box. The algorithm then generates and displays the top five candidate targeting sequences in the "output sequence" box (**Figure 1a**). The candidate targeting sequences are of variable length from ~18 to 30 bp and are displayed in the sense orientation.

We first validated this algorithm for TGS of several oncogenes. The promoters of *AKT-1*, *MYC*, *K-RAS*, and *H-RAS* were targeted, and the top candidate sasRNAs were generated (**Figure 1b**). sasRNAs that effectively induce TGS were identified for all of the targeted oncogene promoters with a success rate greater than 50%. Most effective was an sasRNA targeting *AKT-1* (as-A3, **Figure 1c**) and one directed to *MYC* (as-M1, **Figure 1d**). Transcription of *K-RAS* and *H-RAS* was also susceptible to algorithm-designed sasRNAs (**Supplementary Figure S1a,b**).

A mutational analysis of the algorithm

In order to examine the role of the algorithm-generated targeting sequence and of structural parameters in the induction of TGS, we performed a mutational analysis. We chose the functionally active small RNAs as-A3 and as-M1 (**Figure 1c**,**d**) to introduce two kinds of mutations. In the first kind, we changed the sequence of the conserved purine string (**Figure 1b**). These mutations are subsequently referred to as sequence changes (**Figure 2a** and **Supplementary Figure S2a**). In the second kind of mutations, we introduced structural changes that affect the predicted folding of the sasRNA. These mutations are referred to as structural changes (**Figure 2a** and **Supplementary Figure S2a**). One fundamental difference between these two classes of mutations can be found in their predicted overall thermodynamic stability (**Figure 2a** and **Supplementary Figure S2a**). The thermodynamic stability affects the ability of RNA to interact with target molecules. In these experiments, structural changes in the sasRNAs affected their TGS-inducing activity (**Figure 2c** and **Supplementary Figure S2c**, respectively). Changes in the purine string had a small effect on the induction of TGS. However, in the *AKT*-targeted series, the effect of multiple sequence changes appeared additive (**Supplementary Figure S2b**). Unexpectedly, one of the mutations in Myc1 proved functional (**Figure 2b**). This

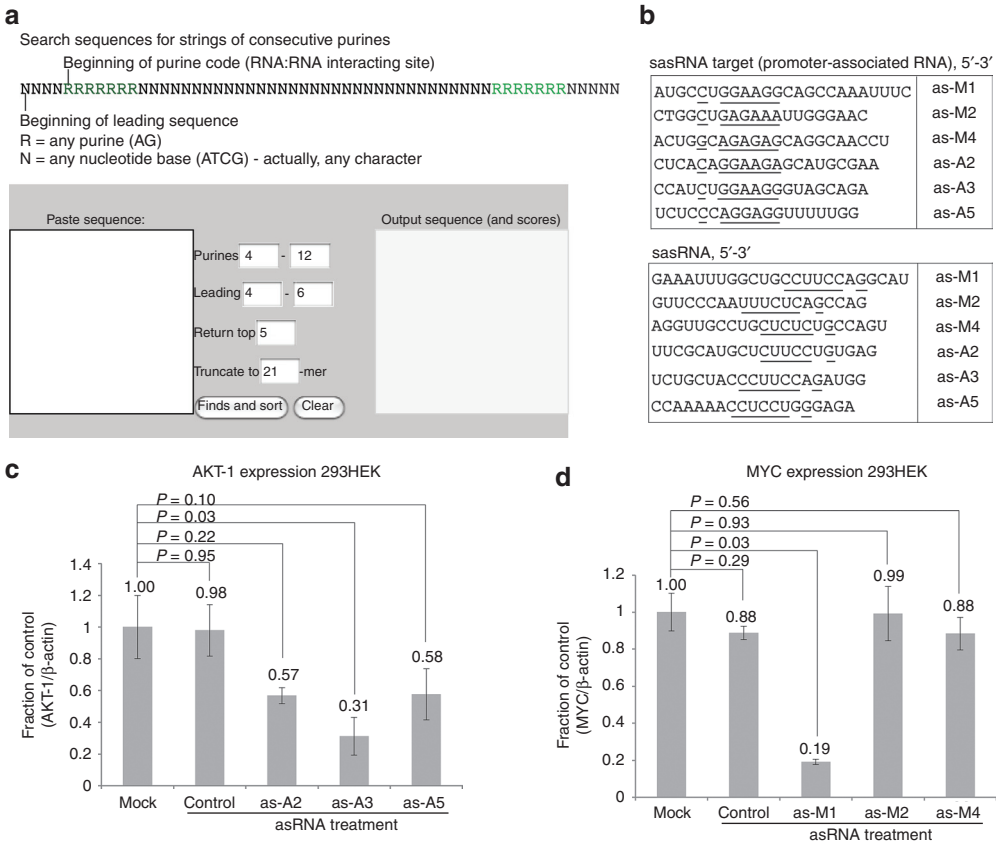

Figure 1 Algorithm-designed small noncoding antisense RNAs (sasRNAs) and transcriptional gene silencing. (**a**) A computational algorithm was designed based on detecting conserved purine strings in promoter targets (http://www.morrislab.unsw.edu.au/TGSalgorithm/Algorithm/purine_finder.html). (**b**) Examples of algorithm-generated sasRNAs that suppress transcription and their promoter target sequences. Critical conserved tracks of nucleotides, the purine code RNA:RNA-interacting sites are underlined. The algorithm was used to generate sasRNAs targeting either (**c**) *AKT* or (**d**) *MYC*. sasRNA constructs are designated as follows: as-M for targeting *MYC* and as-A for targeting *AKT-1*. For **c** and **d**, the averages of triplicate-transfected 293HEK cells are shown with the SEM and *P* values from a paired *t*-test.

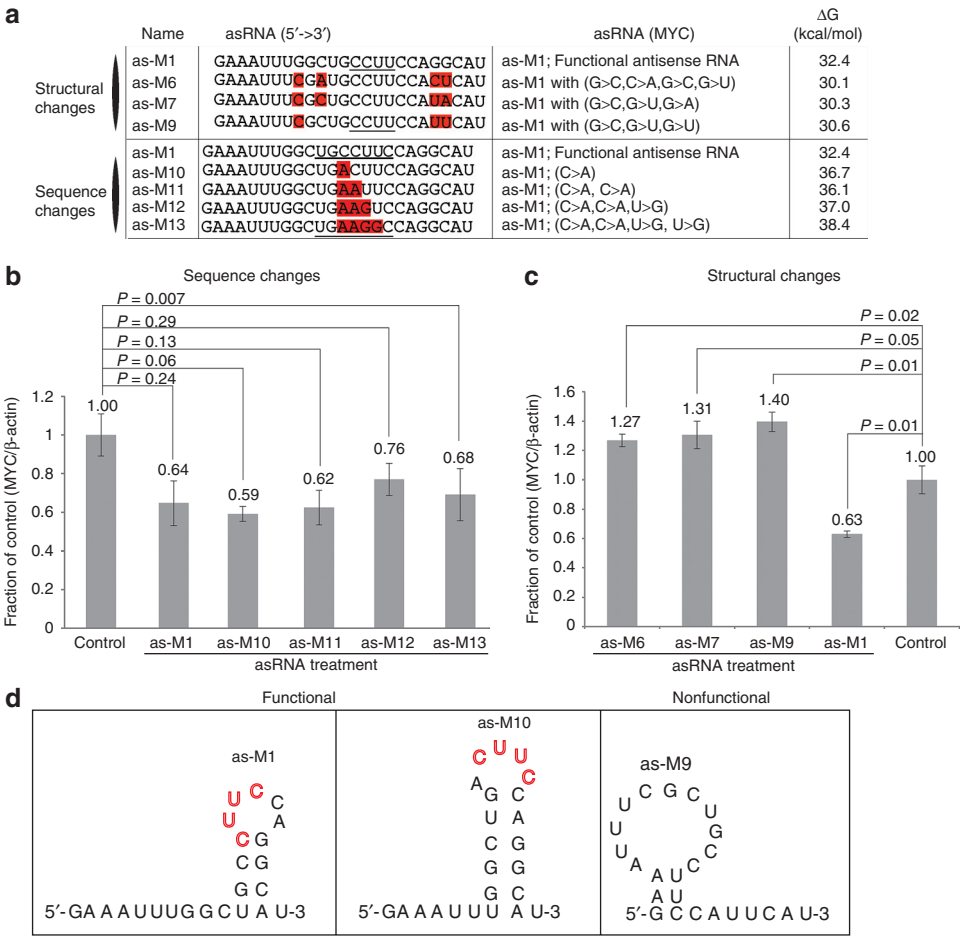

Figure 2 The effects of mutations on small noncoding antisense RNA (sasRNA) function. (**a**) Changes in the sequence and structure of the algorithm-defined sasRNAs for *MYC*. The effects of (**b**) sequence versus (**c**) structural mutations on sasRNA-mediated transcriptional gene silencing of *MYC*. (**d**) M-fold was used to generate the predicted secondary structures for the functional versus nonfunctional sasRNAs. For **b** and **c**, the averages of triplicate-treated cultures are shown with the SEM and *P* values from a paired *t*-test.

sasRNA, termed as-M10, as well as the other active sasRNAs, retained the purine string in the predicted loop of the construct when compared by the M-fold program with inactive constructs (**Figure 2d** and **Supplementary Figure S2d**). This observation may suggest that the sequence GAAC in promoter-associated transcripts, or GUUC in the loop of the small promoter-targeting RNA, is of functional significance in the induction of TGS in human cells (**Figure 1b** and **Supplementary Figure S2e**). However, this is only a supposition, as further studies to delineate this were not carried out. These data also suggest the initial algorithm, while currently useful, can be refined.

Small RNA-induced TGS alters the chromatin architecture of the target locus

Previous studies of TGS induced by either small antisense or long antisense RNA have identified a role of DNA methyltransferase 3A (DNMT3a), histone deacetylase 1 (HDAC-1), and enhancer of zeste 2 (EZH2).[3,6,21–23] In order to determine whether these chromatin-modifying proteins are involved in TGS induced by as-M1, as-M10, or as-A3, we suppressed these chromatin modifiers with RNA interference (**Supplementary Figure S3**). The action of the sasRNAs was abrogated by the suppression of DNMT3a, HDAC-1, or Ezh2, suggesting that these epigenetic remodeling components are involved in the TGS induced by the algorithm-designed sasRNA constructs (**Figure 3a**).

To further confirm that the constructs as-M1 and as-M10 actively induce TGS of the *MYC* gene, we performed quantitative reverse transcription-PCR for the unspliced *MYC* transcripts following treatment of the cells with as-M1 or as-M10. The assessment of unspliced transcripts, an indication of ongoing transcription, suggested that both as-M1 and as-M10 suppress transcription of *MYC* (**Figure 3b**). In these and previous experiments, as-M10 showed higher activity and was, therefore, used in chromatin immunoprecipitation to survey the epigenetic landscape of the *MYC* gene after induction of TGS. These tests revealed a modest gain of H3K27me3 specifically at the as-M10 target locus as well as a reduction in histone 3 at the target locus and various regions around the as-M10-targeted sequence (**Figure 3c**). These observations suggest that the epigenetic architecture of the *MYC* gene was modified by the action of as-M10, similar to previous observations with sasRNAs directing TGS in human cells.[3,21,23–25]

Collectively, these data show that the experimental algorithm is capable of identifying regions in promoters that are susceptible to the induction of sasRNA-mediated TGS. Our

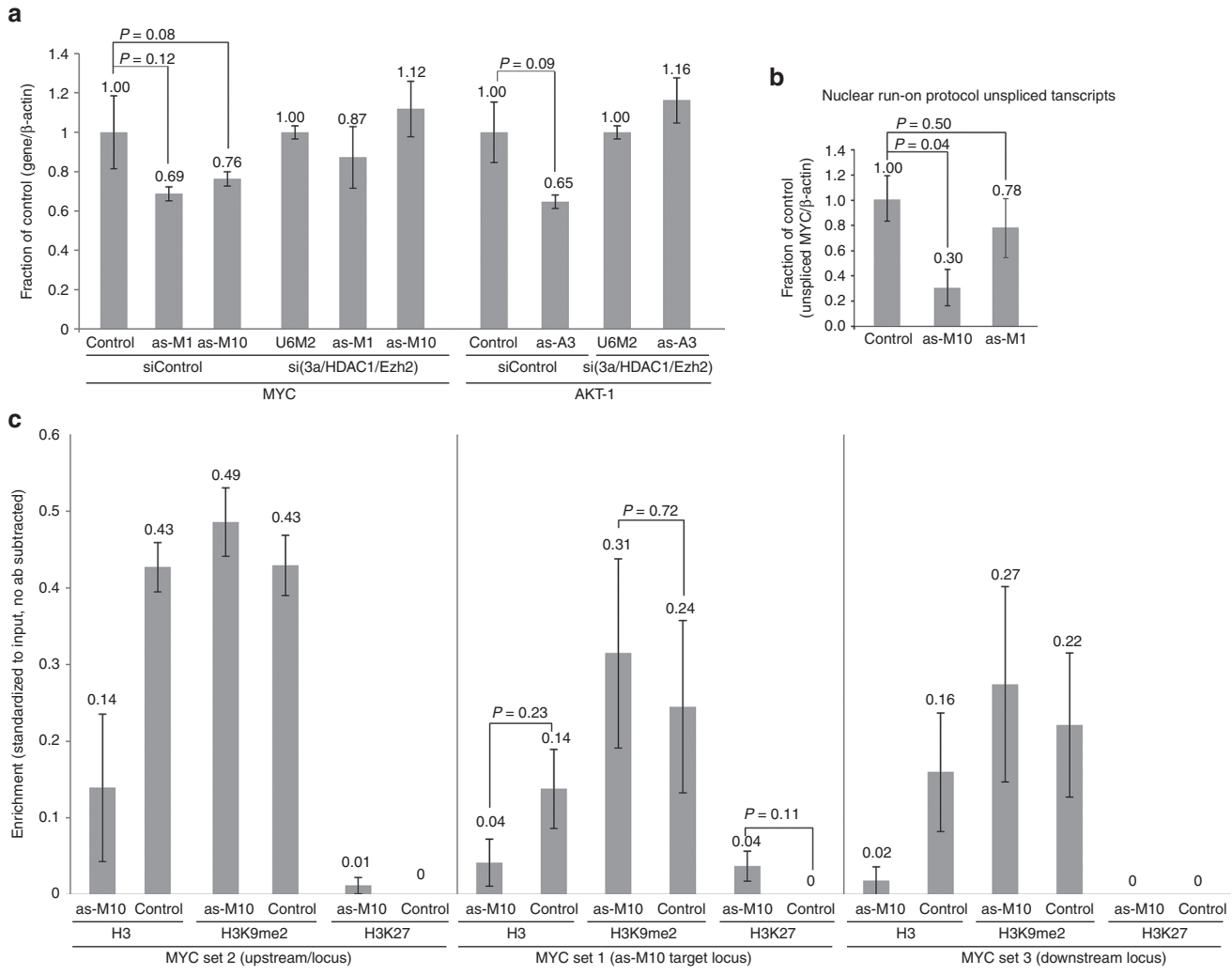

Figure 3 Mechanistic insights from algorithm-generated small noncoding antisense RNAs. (**a**) The effects of DNMT3a, HDAC-1, and Ezh2 expression on the ability of as-M1, as-M10, and as-A3 to induce transcriptional gene silencing were determined. (**b**) The ability of as-M1 and the enhanced functional as-M10, derived from point mutations of as-M1, to direct transcriptional silencing was determined by nuclear run-on and unspliced transcript analysis. (**c**) The effects of as-M10 to epigenetically remodel the *MYC* promoter was assessed by chromatin immunoprecipitation analysis at three loci: an upstream locus, the as-M10 target, and a downstream locus. For **a**–**c**, the averages of triplicate-transfected 293HEK cells are shown with the SEMs and *P* values from a paired *t*-test.

observations also suggest that the retention of the purine string in the loop predicted by the M-fold program may be important for activity.

Targeting an endogenous antisense transcript affects gene expression

Endogenous antisense transcripts (natural antisense transcripts) can modulate the epigenetic states of some genes.[1] In order to explore possible mechanisms for this activity, we used the new algorithm for selecting TGS-active sasRNAs to target the putative promoter of an endogenous antisense transcript, EST AK124265. EST AK124265 was found on the UCSC genome browser to align antisense to the promoter of dual-specificity phosphatase (*DUSP6*) (**Figure 4a**). *DUSP6* is a regulator of extracellular signal-regulated kinase. It is epigenetically silenced in some pancreatic cancers with significant DNA methylation found at the *DUSP6* promoter.[26,27] We surmised that AK124265 is endogenously active in controlling the transcription and epigenetic states of *DUSP6*.

To explore this hypothesis, we first targeted an siRNA construct to a transcribed noncoding RNA sequence of AK124265 and determined the effects on the expression of AK124265 and of *DUSP6* (**Supplementary Figure S4a**). We observed the expected reduction in the level of AK124265 and measured a concomitant increase in the level of *DUSP6* (**Supplementary Figure S4b**). This observation suggested a link between AK124265 and *DUSP6*. We, therefore, attempted to control the levels of AK124265 by targeting the promoter of this natural antisense transcript. We used algorithm-generated sasRNA constructs directed to four regions upstream of the predicted transcriptional start site of AK124265 (**Figure 4a**). These regions showed promoter activity in a luciferase reporter assay (**Figure 4b**). Only one of the algorithm-generated sasRNA constructs, as-AK4, induced elevated expression of *DUSP6* (**Figure 4d**) while

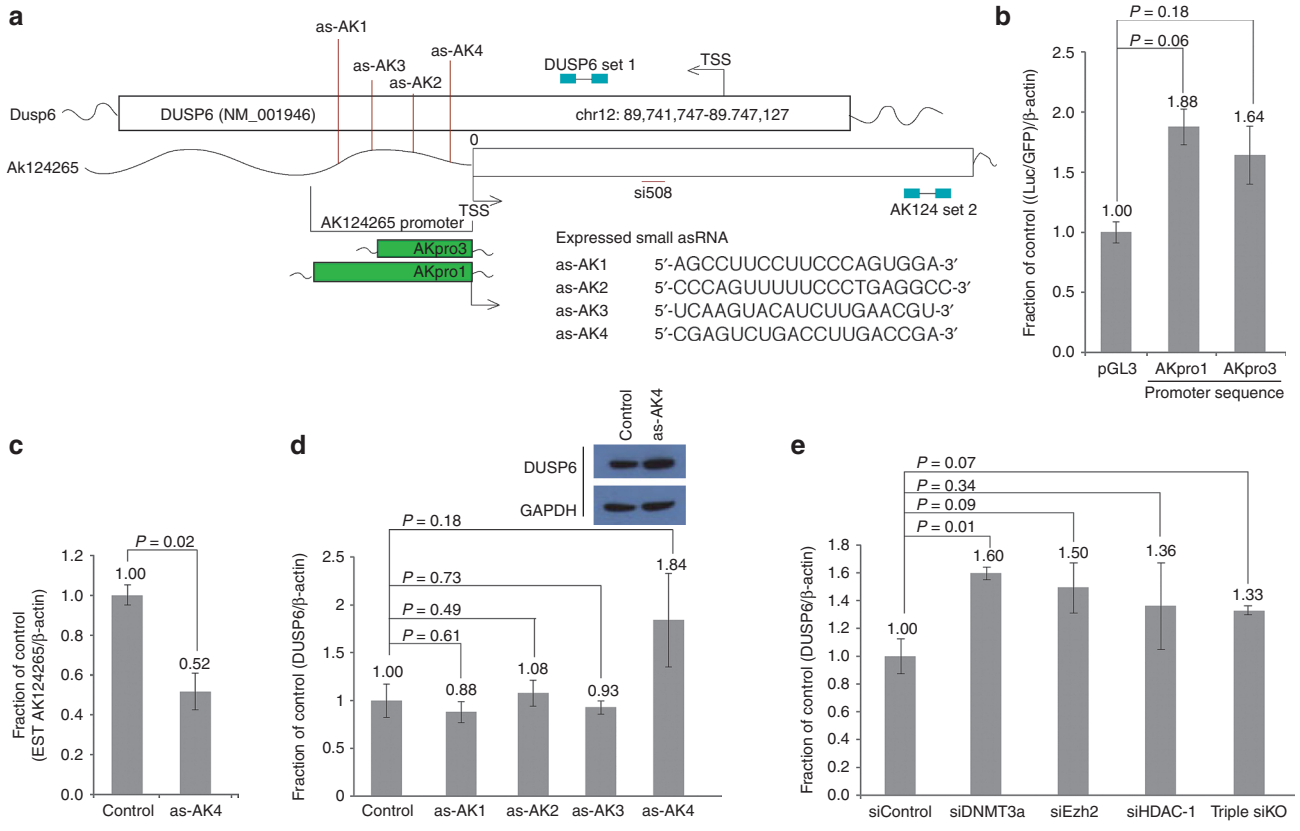

Figure 4 Targeting the promoter of the natural antisense transcript AK124265 with algorithm-guided transcriptional gene silencing. (**a**) A schematic depicting the *DUSP6* gene and the antisense transcript AK124265. Primer sets used in the analysis of this locus as well as fragments of AK124265, which were cloned to assess promoter activity (green boxes), are shown. The small noncoding antisense RNAs (sasRNAs) generated by the algorithm targeting the AK124265 promoter are also shown. (**b**) Assessment of AK124265 promoter activity. The ability of AKpro1 and AKpro3 to express luciferase is shown relative to the parental control pGL3. (**c**) The algorithm-generated sasRNA, as-AK4, suppresses AK124265 expression. (**d**) Algorithm-generated sasRNA, as-AK4, treatment results in the activation of *DUSP6* as determined in quantitative reverse transcription-PCR and western blot analysis (inset). (**e**) Suppression of the chromatin modifiers, DNMT3a, Ezh2, and HDAC-1, leads to enhanced expression of *DUSP6*. For **b**–**e**, the averages of triplicate-transfected 293HEK cells are shown with the SEMs and *P* values from a paired *t*-test.

suppressing AK124265 (**Figure 4c**). The suppression of DNMT3a or DNMT3a, HDAC-1, and Ezh2 in combination also induced marked increases in *DUSP6* (**Figure 4e**), suggesting that *DUSP6* is endogenously regulated by AK124265 by RNA-directed TGS.

Discussion

In this study, we have used algorithm-designed sasRNAs to upregulate and to downregulate transcription from specific genes. Our observations suggest that in human cells, such sasRNAs can affect promoter activity of protein-coding and of noncoding transcripts. Targeting the promoter of a protein-coding transcript can reduce transcription of the gene, whereas targeting of the promoter of the corresponding native antisense transcript results in the elevation of gene expression.

The induction of TGS by siRNAs that target promoter sequences has been documented for several genes (reviewed in ref. 2). Available evidence suggests that this activity involves chromatin-modifying mechanisms. In cell culture, TGS is stable over prolonged time periods due to an epigenetic-based mechanism involving both histone and DNA methylation being directed to the small RNA-targeted promoter loci.[3] The computer algorithm presented in this paper permits the design of RNA constructs that efficiently induce TGS. Important features of this TGS include targeting of promoter sequences at one to three nucleosomes upstream of the dominant transcription start site and the involvement of chromatin-modifying proteins. Constructs designed by the algorithm have proved effective in reducing transcriptional activities from the promoters of *MYC* and of *H-RAS*. The transcriptional activity of *MYC* has been reduced with siRNAs before,[28] but one of these constructs did not target the promoter upstream and probably worked through blockage of RNA polymerase II rather than through epigenetic change.[17] The algorithm described here is empirically derived and incorporates both sequence and structural features. It enhances the success rate of designing sasRNA constructs that are effective in transcriptional regulation but should be considered a "β" version and is open to modification and improvement.

The successful targeting of natural antisense transcripts with small RNAs is less common.[4,5,29] In mammalian cells, antisense transcripts were first discovered by Williams and Fried 1986;[30] another antisense transcript identified early

was that of *MYC*.[31] Today endogenous antisense transcription is recognized as pervasive and as an integral cellular activity.[32–35] The regulatory potential of antisense transcripts was immediately recognized,[36] and their main activity was initially linked to interference with protein translation.[37,38] However, more recent observations suggest an important role of natural antisense transcripts in the regulation of gene transcription by epigenetic modification of chromatin.[4–6,29,39]

We have used the newly developed computer algorithm to target sasRNAs for the first time to a putative promoter of a natural antisense transcript, EST AK124265. This natural antisense transcript spans the promoter of *DUSP6*, a tumor suppressor which is transcribed from the complementary DNA strand. The underlying assumption in targeting the promoter of EST AK124265 is that this antisense transcript functions as a negative regulator of *DUSP6* whose promoter is heavily methylated and epigenetically silenced in pancreatic cancers.[26,27] The transcriptional derepression of *DUSP6* by sasRNA directed at the EST AK124265 promoter provides initial support for this assumption. More detailed studies on the molecular mechanism of this transcriptional activation are currently in progress.

sasRNAs directed to promoter sequences can be used both to inhibit and to activate transcription, depending on the sense of the target DNA strand. In general, the inhibition of active gene function can be readily accomplished. Options besides sasRNA-induced TGS include microRNAs that interfere with translation or direct inhibition of the target protein by small molecules or biomolecules. In contrast, restoration of lost function is a much greater challenge. Natural antisense transcripts may be widely involved in epigenetic gene silencing (discussed in refs. 13,40), and targeting the promoters of such transcripts could provide a valuable technique for the reactivation of silenced genes. The algorithm presented here will facilitate this novel approach to the restoration of gene expression.

Materials and methods

Cell culture and sasRNA validation studies. Human 293HEK and CFPAC cell lines were transfected with the sasRNA pU6M2-based expressing plasmids[41] at a ratio of ~1 μg/10^6 cells. The "Control" used in this study was the U6M2 plasmid without any sasRNA expressed. Transfection of 293HEK and CFPAC cells was carried out using Lipofectamine 2000. Transfected cultures were assessed 72 hours post-transfection and gene expression was determined by quantitative reverse transcription-PCR.

Small RNA-expressing vectors. sasRNAs were cloned into the pU6M2 expression vector (**Supplementary Figure S5**). For each expressed small RNA, two oligonucleotides were designed containing a 5′BglI-II and 3′ Kpn-I restriction sites (**Supplementary Table S3**). The oligonucleotides were then treated with T4 polynucleotide kinase and heat-inactivated at 65 °C for 20 minutes. The oligonucleotides were mixed, incubated in a boiling water bath for 1 minute, and then allowed to slowly cool to room temperature. The oligonucleotides were then diluted 1:25 in water, digested, and ligated into the pU6M2 plasmid that had been treated the same way. Positive clones were collected (Qiagen QIAprep Miniprep on the Qiacube; Qiagen, Valenica, CA) and screened by PCR (Kapa Biosystems, Woburn, MA) using p008 and p080 (**Supplementary Table S3**) and sequenced.

RNA interference targeting DUSP6, AK124265, Ezh2, DNMT3a, and HDAC-1. siRNAs were generated to target either *DUSP6* NM_022652 (siDUSP-375) or EST AK124265 (siAK124265-508) using the Ambion Silencer kit according to the manufacture's instructions (Life Technologies, Carlsbad, CA) (**Supplementary Table S3**). The siRNAs were transfected into CFPAC cells (100 nmol/l) using Lipofectamine 2000 (L2K) according to the manufacturer's instructions (Life Technologies).

siRNAs for targeting HDAC-1, Ezh2, and DNMT3a were generated from Qiagen, Ezh2 (catalog no. s102665166), HDAC-1 (catalog no. s102663472), and DNMT3a (catalog no. s102665278). The siRNAs were transfected (50 nmol/l) using RNAiMax (Life Technologies).

Western blots. Western blot analysis was carried out on CFPAC cells 72 hours following transfection with either the as-AK4 or control U6M2 plasmids. The *DUSP6* antibody used was Antibody: ab54940 (Abcam, Cambridge, UK), and the glyceraldehyde-3-phosphate dehydrogenase antibody used was anti-GAPDH (mouse monoclonal, 1:5,000 dilution, Millipore: CB1001; Millipore, Billerica, MA) following previously described protocols.[42]

RNA isolation and gene expression profiling. The transiently transfected cells were collected at predetermined time points and RNA was extracted (Qiacube Qiagen; Qiagen). The cellular RNA was DNAse treated (Ambion Turbo DNAse; Life Technologies), and cDNA was generated. The resultant cDNA libraries from each sample were then subjected to quantitative PCR using the Eppendorf real-time cycler and gene-specific primers (**Supplementary Table S3**). To assess changes at the transcriptional level by nuclear run-on, the previously reported technique for nuclear run-on was adopted[3,29] with the minor exception that primers for unspliced forms of *MYC* were utilized to measure gene expression (**Supplementary Table S3**).

NAT promoter analysis. To determine whether genomic regions just upstream of the reported NAT EST AK124265 are functional promoter elements, PCR of various regions was performed. The AK124265 promoter, specifically AKPro1, was cloned using the AKproF1 and AKproR1, whereas the AKPro3 was cloned using AKproF3 and AKproR1 primers (**Supplementary Table S3**). The resultant PCR products, Xho-I and Mlu-I, were digested and ligated into the equally treated pGL3 Enhancer Vector (Promega, Madison, WI). Luciferase expression was determined by quantitative reverse transcription-PCR following transfection into 293HEK cells.

Algorithm. Characters typed or pasted into the input box will be dynamically colored to indicate whether they are purines, pyrimidines, other International Union of Pure and Applied Chemistry nucleotide characters, comments, or invalid characters. The comments and invalid characters will be stripped out prior to processing. The algorithm can be found at: http://www.morrislab.unsw.edu.au/TGSalgorithm/Algorithm/purine_finder.html and is freely available to use by academic investigators. The algorithm specifically looks for purine

strings from 4 to 12 bp of the particular promoter-targeted locus, based on observations with those small noncoding RNAs that have been observed previously to direct TGS in human cells (contrast **Supplementary Table S1 with Supplementary Table S2**). Those promoter-targeted loci which are utilized in the algorithm are generally one to three nucleosomes upstream of the dominant transcription start site as these regions tend to favor the involvement of chromatin-modifying proteins.[2] The resulting sasRNA target sites should be assessed using BLAT in the UCSC genome browser to confirm whether targeting is specific to the intended target loci and the other loci are not secondarily targeted.

Supplementary material

Figure S1. Algorithm-generated small antisense RNA target validation.
Figure S2. Point mutation analysis of sasRNAs targeted to the *AKT-1* promoter reveals a structural aspect to the algorithm-designed functional asRNAs.
Figure S3. Suppression of epigenetic regulatory factors.
Figure S4. siRNA targeting *DUSP6* and AK124265 in CF-PAC cells.
Figure S5. The U6M2 insertion for generating asRNAs.
Table S1. Functional TGS targets.
Table S2. Nonfunctional TGS targets.
Table S3. Oligonucleotides used in this study.

Acknowledgments. Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award number R01AI084406 and P01AI099783-01 (K.V.M.) and by the National Institute of Cancer of the National Institutes of Health under award numbers CA151574 and RO1 CA153124 (P.K.V.). A provisional patent filing for the algorithm has been initiated as TSRI Case 1538.0/TSR1634P. This is TSRI manuscript 21902. The authors declared no conflict of interest.

1. Morris, KV (2011). The emerging role of RNA in the regulation of gene transcription in human cells. *Semin Cell Dev Biol* **22**: 351–358.
2. Morris, KV (2009). RNA-directed transcriptional gene silencing and activation in human cells. *Oligonucleotides* **19**: 299–306.
3. Hawkins, PG, Santoso, S, Adams, C, Anest, V and Morris, KV (2009). Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells. *Nucleic Acids Res* **37**: 2984–2995.
4. Morris, KV, Santoso, S, Turner, AM, Pastori, C and Hawkins, PG (2008). Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet* **4**: e1000258.
5. Yu, W, Gius, D, Onyango, P, Muldoon-Jacobs, K, Karp, J, Feinberg, AP *et al.* (2008). Epigenetic silencing of tumour suppressor gene *p15* by its antisense RNA. *Nature* **451**: 202–206.
6. Johnsson, P, Ackley, A, Vidarsdottir, L, Lui, WO, Corcoran, M, Grandér, D *et al.* (2013). A pseudogene long-noncoding-RNA network regulates *PTEN* transcription and translation in human cells. *Nat Struct Mol Biol* **20**: 440–446.
7. Alimonti, A, Carracedo, A, Clohessy, JG, Trotman, LC, Nardella, C, Egia, A *et al.* (2010). Subtle variations in Pten dose determine cancer susceptibility. *Nat Genet* **42**: 454–458.
8. Khavari, DA, Sen, GL and Rinn, JL (2010). DNA methylation and epigenetic control of cellular differentiation. *Cell Cycle* **9**: 3880–3883.
9. Bachman, KE, Park, BH, Rhee, I, Rajagopalan, H, Herman, JG, Baylin, SB *et al.* (2003). Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. *Cancer Cell* **3**: 89–95.
10. Esteller, M (2007). Epigenetics provides a new generation of oncogenes and tumour-suppressor genes. *Br J Cancer* **96** (suppl.): R26–R30.
11. Furukawa, T and Horii, A (2006). Mechanistic insights for pancreatic carcinogenesis. *Nippon Rinsho* **64** (suppl. 1): 27–31.
12. Morris, KV (2009). Non-coding RNAs, epigenetic memory and the passage of information to progeny. *RNA Biol* **6**: 242–247.
13. Morris, KV (2009). Long antisense non-coding RNAs function to direct epigenetic complexes that regulate transcription in human cells. *Epigenetics* **4**: 296–301.
14. Heale, BS, Soifer, HS, Bowers, C and Rossi JJ (2005). siRNA target site secondary structure predictions using local stable substructures. *Nucleic Acids Res* **33**: e30.
15. Naito, Y, Yamada, T, Ui-Tei, K, Morishita, S and Saigo, K (2004). siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference. *Nucleic Acids Res* **32**: W124.
16. Han, J, Kim, D and Morris, KV (2007). Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proc Natl Acad Sci USA* **104**: 12422–12427.
17. Napoli, S, Pastori, C, Magistri, M, Carbone, GM and Catapano, CV (2009). Promoter-specific transcriptional interference and c-myc gene silencing by siRNAs in human cells. *EMBO J* **28**: 1708–1719.
18. Ross, JP, Suetake, I, Tajima, S and Molloy, PL (2010). Recombinant mammalian DNA methyltransferase activity on model transcriptional gene silencing short RNA-DNA heteroduplex substrates. *Biochem J* **432**: 323–332.
19. Chu, C, Qu, K, Zhong, FL, Artandi, SE and Chang, HY (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell* **44**: 667–678.
20. Kretz, M, Siprashvili, Z, Chu, C, Webster, DE, Zehnder, A, Qu, K *et al.* (2013). Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493**: 231–235.
21. Suzuki, K, Juelich, T, Lim, H, Ishida, T, Watanebe, T, Cooper, DA *et al.* (2008). Closed chromatin architecture is induced by an RNA duplex targeting the HIV-1 promoter region. *J Biol Chem* **283**: 23353–23363.
22. Turner, AM, De La Cruz, J and Morris, KV (2009). Mobilization-competent lentiviral vector-mediated sustained transcriptional modulation of HIV-1 expression. *Mol Ther* **17**: 360–368.
23. Weinberg, MS, Villeneuve, LM, Ehsani, A, Amarzguioui, M, Aagaard, L, Chen, ZX *et al.* (2006). The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells. *RNA* **12**: 256–262.
24. Morris, KV, Chan, SW, Jacobsen, SE and Looney, DJ (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* **305**: 1289–1292.
25. Turner, AM, Ackley, AM, Matrone, MA and Morris, KV (2012). Characterization of an HIV-targeted transcriptional gene-silencing RNA in primary cells. *Hum Gene Ther* **23**: 473–483.
26. Furukawa, T, Sunamura, M, Motoi, F, Matsuno, S and Horii, A (2003). Potential tumor suppressive pathway involving DUSP6/MKP-3 in pancreatic cancer. *Am J Pathol* **162**: 1807–1815.
27. Xu, S, Furukawa, T, Kanai, N, Sunamura, M and Horii, A (2005). Abrogation of DUSP6 by hypermethylation in human pancreatic cancer. *J Hum Genet* **50**: 159–167.
28. Mehndiratta, M, Palanichamy, JK, Pal, A, Bhagat, M, Singh, A, Sinha, S *et al.* (2011). CpG hypermethylation of the C-myc promoter by dsRNA results in growth suppression. *Mol Pharm* **8**: 2302–2309.
29. Hawkins, PG and Morris, KV (2010). Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. *Transcription* **1**: 165–175.
30. Williams, T and Fried, M (1986). A mouse locus at which transcription from both DNA strands produces mRNAs complementary at their 3' ends. *Nature* **322**: 275–279.
31. Spicer, DB and Sonenshein, GE (1992). An antisense promoter of the murine c-myc gene is localized within intron 2. *Mol Cell Biol* **12**: 1324–1329.
32. Clark, MB, Amaral, PP, Schlesinger, FJ, Dinger, ME, Taft, RJ, Rinn, JL *et al.* (2011). The reality of pervasive transcription. *PLoS Biol* **9**: e1000625; discussion e1001102.
33. Mattick, JS and Makunin, IV (2006). Non-coding RNA. *Hum Mol Genet* **15** Spec No 1: R17–R29.
34. Numata, K, Osada, Y, Okada, Y, Saito, R, Hiraiwa, N, Nakaoka, H *et al.* (2009). Identification of novel endogenous antisense transcripts by DNA microarray analysis targeting complementary strand of annotated genes. *BMC Genomics* **10**: 392.
35. Lapidot, M and Pilpel, Y (2006). Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO Rep* **7**: 1216–1222.
36. Kumar, M and Carmichael, GG (1998). Antisense RNA: function and fate of duplex RNA in cells of higher eukaryotes. *Microbiol Mol Biol Rev* **62**: 1415–1434.
37. Knee, R and Murphy, PR (1997). Regulation of gene expression by natural antisense RNA transcripts. *Neurochem Int* **31**: 379–392.
38. Stolt, P and Zillig, W (1993). Antisense RNA mediates transcriptional processing in an archaebacterium, indicating a novel kind of RNase activity. *Mol Microbiol* **7**: 875–882.
39. Su, WY, Xiong, H and Fang, JY (2010). Natural antisense transcripts regulate gene expression in an epigenetic manner. *Biochem Biophys Res Commun* **396**: 177–181.
40. Morris, KV and Vogt, PK (2010). Long antisense non-coding RNAs and their role in transcription and oncogenesis. *Cell Cycle* **9**: 2544–2547.
41. Amarzguioui, M, Rossi, JJ and Kim, D (2005). Approaches for chemically synthesized siRNA and vector-mediated RNAi. *FEBS Lett* **579**: 5974–5981.
42. Taft, RJ, Hawkins, PG, Mattick, JS and Morris, KV (2011). The relationship between transcription initiation RNAs and CCCTC-binding factor (CTCF) localization. *Epigenetics Chromatin* **4**: 13.

Supplementary Information accompanies this paper on the Molecular Therapy–Nucleic Acids website (http://www.nature.com/mtna)