Arteriosclerosis, Thrombosis, and Vascular Biology

JOURNAL OF THE AMERICAN HEART ASSOCIATION

Identification and Initial Functional Characterization of a Human Vascular Cell–Enriched Long Noncoding RNA

Robert D. Bell, Xiaochun Long, Mingyan Lin, Jan H. Bergmann, Vivek Nanda, Sarah L. Cowan, Qian Zhou, Yu Han, David L. Spector, Deyou Zheng and Joseph M. Miano

Arterioscler Thromb Vasc Biol. published online February 27, 2014;
Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1079-5642. Online ISSN: 1524-4636

Original Article

Identification and Initial Functional Characterization of a Human Vascular Cell–Enriched Long Noncoding RNA

Robert D. Bell,* Xiaochun Long,* Mingyan Lin,* Jan H. Bergmann, Vivek Nanda, Sarah L. Cowan, Qian Zhou, Yu Han, David L. Spector, Deyou Zheng, Joseph M. Miano

Objective—Long noncoding RNAs (lncRNAs) represent a rapidly growing class of RNA genes with functions related primarily to transcriptional and post-transcriptional control of gene expression. There is a paucity of information about lncRNA expression and function in human vascular cells. Thus, we set out to identify novel lncRNA genes in human vascular smooth muscle cells and to gain insight into their role in the control of smooth muscle cell phenotypes.

Approach and Results—RNA sequencing of human coronary artery smooth muscle cells revealed 31 unannotated lncRNAs, including a vascular cell–enriched lncRNA (smooth muscle and endothelial cell–enriched migration/differentiation-associated long noncoding RNA [*SENCR*]). Strand-specific reverse transcription polymerase chain reaction (PCR) and rapid amplification of cDNA ends indicate that *SENCR* is transcribed antisense from the 5′ end of the *FLI1* gene and exists as 2 splice variants. RNA fluorescence in situ hybridization and biochemical fractionation studies demonstrate *SENCR* is a cytoplasmic lncRNA. Consistent with this observation, knockdown studies reveal little to no *cis*-acting effect of *SENCR* on *FLI1* or neighboring gene expression. RNA-sequencing experiments in smooth muscle cells after *SENCR* knockdown disclose decreased expression of Myocardin and numerous smooth muscle contractile genes, whereas several promigratory genes are increased. Reverse transcription PCR and Western blotting experiments validate several differentially expressed genes after *SENCR* knockdown. Loss-of-function studies in scratch wound and Boyden chamber assays support *SENCR* as an inhibitor of smooth muscle cell migration.

Conclusions—*SENCR* is a new vascular cell–enriched, cytoplasmic lncRNA that seems to stabilize the smooth muscle cell contractile phenotype. **(*Arterioscler Thromb Vasc Biol*. 2014;34:00-00.)**

Key Words: cell migration ■ endothelial cells ■ myocytes, smooth muscle ■ RNA, long noncoding ■ RNA sequence

Although the human genome is 30× larger than that of *Caenorhabditis elegans*, each species is endowed with a similar number of protein-coding genes, a fact seemingly in support of an abundance of junk DNA within our genome.[1] Two major discoveries during the past 10 years challenge this decades-old concept. First, genome-wide RNA expression studies show widespread transcription across the mouse and human genomes with roughly equal amounts of polyadenylated and nonpolyadenylated RNA.[2–7] Second, the combined efforts of the Encyclopedia of DNA Elements Consortium and many other laboratories have revealed the existence of millions of codes that punctuate the human genome, most notably codes for transcription factor binding.[8–12] These findings, coupled with the notion that much of the human genome is functional with 50% to 90% comprising transcribed sequences,[13,14] debunk the concept of junk DNA and point to a genome replete with information essential for human life.

Much of the noncoding RNA (ncRNA) in a cell functions to orchestrate basic translation (transfer and ribosomal RNA); however, 2 broad classes of ncRNA expanded greatly at the turn of the millennium, primarily as a result of large-scale transcriptomics projects.[2,3,15] These ncRNAs are classified subjectively as either short (processed transcript length <200 nucleotides) or long (processed transcript length >200 nucleotides). Short ncRNAs include small nucleolar RNA and their derivatives that act as guide RNAs to modify ribosomal and transfer RNAs,[16] as well as microRNA, small interfering RNA, and PIWI-interacting RNA that use Argonaute proteins to mediate endonucleolytic cleavage of target RNAs.[17]

Received on: January 11, 2014; final version accepted on: February 12, 2014.

From the Aab Cardiovascular Research Institute, University of Rochester School of Medicine and Dentistry, Rochester, NY (R.D.B., X.L., V.N., S.L.C., Q.Z., Y.H., J.M.M.); Department of Genetics (M.L., D.Z.) and Departments of Neurology and Neuroscience (D.Z.), Albert Einstein College of Medicine, Bronx, NY; and Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (J.H.B., D.L.S.).

*These authors share equal authorship.

The online-only Data Supplement is available with this article at http://atvb.ahajournals.org/lookup/suppl/doi:10.1161/ATVBAHA.114.303240/-/DC1.

Correspondence to Joseph M. Miano, Aab Cardiovascular Research Institute, University of Rochester School of Medicine and Dentistry, 601 Elmwood Ave, Rochester, NY 14642. E-mail j.m.miano@rochester.edu or Deyou Zheng, Department of Neurology, Albert Einstein College of Medicine, 1300 Morris Park Ave, Bronx, NY 10461. E-mail deyou.zheng@einstein.yu.edu

Nonstandard Abbreviations and Acronyms

EC	endothelial cell
FLI1	Friend leukemia virus integration 1
HCASMC	human coronary artery smooth muscle cell(s)
HUVEC	human umbilical vein endothelial cell(s)
lincRNA	long intervening noncoding RNA
lncRNA	long noncoding RNA
MYOCD	myocardin
RT-PCR	reverse transcription polymerase chain reaction
SENCR	smooth muscle and endothelial cell enriched migration/differentiation-associated long noncoding RNA
SMC	smooth muscle cell

Long ncRNAs (lncRNAs) function in a myriad of biological processes and may be classified loosely based on their physical location in the genome. Long intervening ncRNAs (lincRNAs) are a subclass of lncRNAs found between 2 transcription units, and they exhibit similar active chromatin signatures as those found around protein-coding genes.[18–20] LincRNAs may display tissue-specific patterns of expression and function principally as scaffold or guide RNAs that facilitate chromatin remodeling in *cis* or *trans* to directly influence gene transcription (nuclear lincRNAs) or effect changes in mRNA stability/protein translation (cytoplasmic lincRNAs).[20–22] Examples of lincRNAs include the abundantly expressed *MALAT1* that functions in processing of mRNAs[23] and the epidermal prodifferentiating *TINCR*.[24] A recent report defined long intervening ncRNAs (≤700 kb) whose expression correlates with malignancy; these transcripts may encompass previously annotated lincRNAs.[25] LincRNAs may also overlap transcriptional enhancers to effect *cis*-mediated changes in gene expression.[26,27]

Intragenic lncRNAs represent another subclass of RNA genes that reside on the sense or antisense strand of an overlapping gene. Sense lncRNAs have been reported only sporadically,[28] although a recent report contends there exists a large number of ill-defined sense ncRNAs within introns.[29] Antisense lncRNAs occur in a significant number of protein-coding genes and may overlap the 5′ or 3′ end of a gene, occur entirely within an intron, or overlap multiple exons.[30–32] Antisense lncRNAs whose exons overlap protein-coding (or ncRNA) exons are known as natural antisense transcripts and these can function in *cis* or *trans* to negatively or positively regulate gene expression through RNA interactions with chromatin remodeling factors.[33] Examples of natural antisense transcripts include the X chromosome inactivating *XIST*[34] and the cell cycle regulator *ANRIL*.[35] Some processed antisense lncRNAs do not overlap sense exons and thus may have unexpected functions (below). The number of human lncRNAs is soaring with the current catalogue of LNCipedia[36] listing >32 000 (http://www.lncipedia.org/), a number that exceeds all protein-coding genes. Thus, lncRNAs embody a rapidly growing class of genes with functions related primarily to the regulation of gene/protein expression.

Cellular differentiation requires the coordinated activation of unique gene sets through transcription factors in association with cofactors over discrete *cis* elements. For example, vascular smooth muscle cell (SMC) differentiation is chiefly a function of ubiquitously expressed serum response factor[37] binding a cardiovascular-restricted cofactor called myocardin (MYOCD)[38] over CArG elements located in the proximal promoter region of many SMC-associated genes.[39] Similarly, endothelial cell (EC) differentiation proceeds, in part, through the FOXC2[40] and ETV2[41] transcription factors binding a composite *cis* element, the FOX-ETS motif, found in promoter/enhancer sequences of several EC-specific genes.[42] Normal differentiated properties of SMC and EC further require fine tuning of gene expression through the action of microRNAs.[43] Because lncRNAs are prevalent and play key roles in modulating gene expression,[44] they too may have functions linked to vascular cell phenotype. Little is known, however, about the expression or function of lncRNAs in vascular cells,[45–49] and there is nothing known about human-specific, vascular cell–selective lncRNAs. Accordingly, we performed RNA-seq in human coronary artery smooth muscle cell (HCASMC) as a first step toward understanding the potential role of lncRNAs in human SMC phenotypic control. Here, we report on the identification of 31 lncRNAs, including 1 named *SENCR* (for smooth muscle and endothelial cell–enriched migration/differentiation-associated long noncoding RNA). We have characterized the expression, splicing, and localization of *SENCR* and have identified unique gene signatures on its knockdown in SMC. *SENCR* seems to play a role in maintaining the normal SMC differentiated state as its attenuated expression leads to reduced *MYOCD* and contractile gene expression with elevations in migratory genes that foster a hyper-motile state. This report outlines the first foray into lncRNA discovery in human vascular cells and establishes a foundation for further inquiry into *SENCR* biology, as well as the identification, expression, and function of other human vascular-selective lncRNAs under normal and pathological cell states.

Materials and Methods

Materials and Methods are available in the online-only Supplement.

Results

Identification and Validation of lncRNAs in HCASMC

We have developed a rigorous workflow for the identification and study of lncRNAs in primary-derived HCASMC using RNA-seq methodology (Figure I in the online-only Data Supplement). A total of 79.41% of filtered reads could be aligned to the human reference genome. Thirty-one lncRNAs met our strict inclusion criteria (Methods in the online-only Data Supplement) with the majority (22/31) falling into the lincRNA subclass (Table II in the online-only Data supplement). Conventional reverse transcription polymerase chain reaction (RT-PCR) showed detectable expression of 21 of 31 lncRNAs in a panel of human cell types, including HCASMC and human umbilical vein EC (HUVEC; Figure 1A). Sequence analysis of the PCR products confirmed the identity of each lncRNA (not shown). The majority of HCASMC lncRNAs are distributed widely across human tissues with several detected in dated human

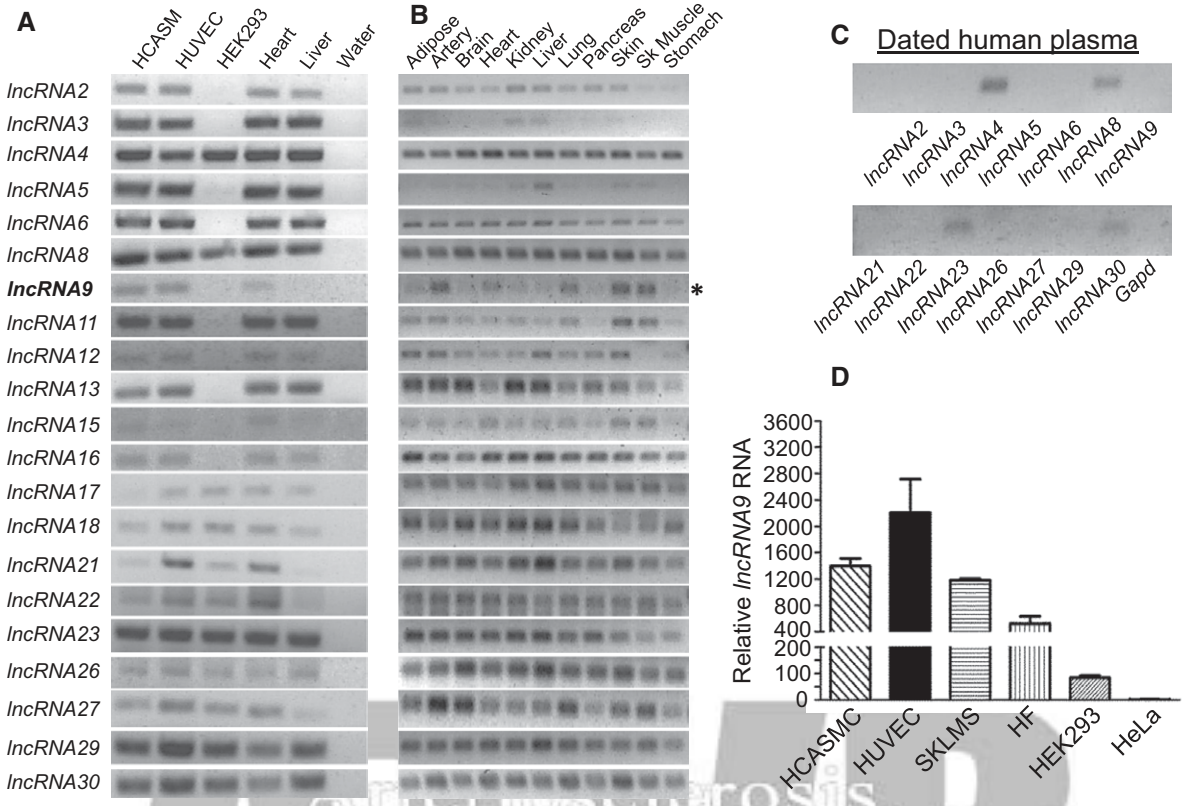


Figure 1. Validation of Long noncoding RNA (lncRNA) expression in human cells and tissues. Reverse transcription polymerase chain reaction (RT-PCR) analysis of 21 lncRNAs (arbitrarily numbered) in indicated human cells (**A**) and tissues (**B**). Bold *lncRNA9* and asterisk denote *SENCR*. **C**, RT-PCR of indicated lncRNAs in dated human plasma. All reactions were done using the same PCR parameters. **D**, Quantitative RT-PCR of *lncRNA9* in the indicated human cell types. HCASMC indicates human coronary artery smooth muscle cell; HF, human fibroblasts; HUVEC, human umbilical vein endothelial cell; and SKLMS, uterine leiomyosarcoma cell line.

plasma (Figure 1B and 1C). One of the lncRNAs (*lncRNA9*) exhibited a selective pattern of expression in cell lines (Figure 1A and 1D) and human tissues (Figure 1A and 1B). We refer to this lncRNA as *SENCR* because of its enriched expression in both smooth muscle and ECs (Figure 1A and 1D) and its proposed function (below).

SENCR Is a Vascular Cell–Selective Antisense lncRNA

RNA-seq alignment, 5′ RACE, and RT-PCR with oligo-dT and strand-specific primers established that *SENCR* comprises 3 exons and is transcribed in the antisense orientation from within the first intron of Friend leukemia virus integration 1 (*FLI1*), an important transcription factor programming EC and blood cell formation[50] (Figure 2A). There is no overlap between *SENCR* and *FLI1* exonic sequences, indicating that *SENCR* is not a natural antisense transcript[33] (Figure 2A). The longest open reading frame flanked by start and stop codons is 61 amino acids; however, analysis of this and other predicted open reading frames in *SENCR* failed to reveal any known protein-coding domains, suggesting that this transcript has no or low protein-coding potential (not shown). Primers to exons 1 and 3 of *SENCR* showed the presence of 2 distinct PCR products (Figure 2B). Sequence analysis confirmed these products as full length (*SENCR_V1*) and an alternatively spliced variant (*SENCR_V2*) of the *SENCR* gene (Figure 2A and 2B). These sequences have been deposited in GenBank under accession numbers KF806591 and KF806590, respectively. We used specific primer pairs to examine *SENCR* isoform expression in a panel of human tissues and cell lines. Results showed *SENCR_V1* to be more broadly expressed than *SENCR_V2* (Figure 2C and 2D). In general, there was coincident expression of *SENCR* with *FLI1*, suggesting that these transcripts may be under similar transcriptional control processes (Figure 2E and 2F). Quantitative RT-PCR analysis suggested the *FLI1* transcript to have higher expression than *SENCR* (Figure II in the online-only Data Supplement).

Exon 1 of *FLI1* shows high conservation across 46 mammalian species; however, much less conservation exists across the 3 exons of *SENCR* (Figure 3), consistent with the fact that no orthologous *SENCR* transcripts have yet been found outside human/chimp lineages. Interestingly, exons 2 and 3 of *SENCR* harbor single nucleotide polymorphisms, suggesting potential deleterious effects on *SENCR* function (Figure 3). Analysis of Encyclopedia of DNA Elements data on the UCSC Genome Browser (http://genome.ucsc.edu/) supports the enriched expression of *SENCR* in HUVEC with lower levels in other cell types. Further, there is a prominent HUVEC-associated H3K4me3 mark near exon 1 of *SENCR*, suggesting the presence of an active promoter (Figure 3). As a first step toward delineating

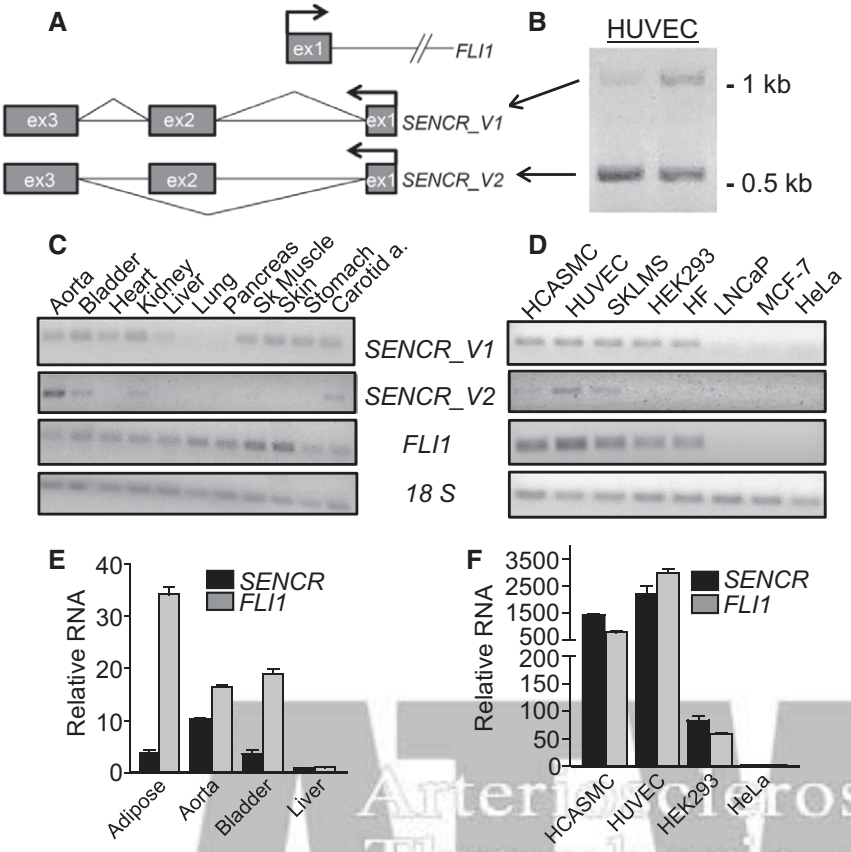


Figure 2. *SENCR* gene structure and isoform expression. **A**, Schematic of *SENCR* and *FLI1* (partial) gene loci. Arrows denote the transcription start sites and bent lines in *SENCR* indicate splicing patterns. **B**, Reverse transcription polymerase chain reaction (RT-PCR) of *SENCR* with primers to exons 1 and 3 showing the presence of 2 transcripts reflecting full length (V1) and alternately spliced (V2) *SENCR*. RT-PCR of 2 *SENCR* isoforms and *FLI1* in various human tissues (**C**) and cell lines (**D**). Quantitative RT-PCR of *SENCR* and *FLI1* in select human tissues (**E**) and cell lines (**F**). Bars here and below represent the SD of 1 experiment with 3 biological replicates. All expression data here and below represent ≥2 (more typically multiple) independent studies performed by >1 author. Unless indicated otherwise, *SENCR* expression here and below reflects both isoforms using primers to a common exon. HCASMC indicates human coronary artery smooth muscle cell; HF, human fibroblasts; HUVEC, human umbilical vein endothelial cell; and SKLMS, uterine leiomyosarcoma cell line.

SENCR transcription, we cloned and tested several luciferase reporter constructs. Luciferase assays showed little to no detectable *SENCR* promoter activity in HUVEC unless sequences encompassing the 5′ *FLI1* promoter region were included, although even these reporters showed much lower activity than a control promoter construct (not shown). Collectively, these results define an alternatively spliced, vascular cell–enriched antisense lncRNA that overlaps the 5′ end of the *FLI1* transcription factor yet, in its mature form, does not harbor exonic sequences that could undergo Watson–Crick base-pairing with corresponding exonic sequences in *FLI1*.

SENCR Is a Cytoplasmic lncRNA

Quantitative RT-PCR showed *SENCR* RNA to be most abundant in HUVEC with undetectable transcripts in HeLa cells (Figure 1D). We used high-resolution RNA fluorescence in situ hybridization[51] in these 2 cells types to unambiguously discern the intracellular compartment where *SENCR* transcripts reside. Consistent with quantitative RT-PCR, no *SENCR* transcripts were seen in individual HeLa cells (Figure 4A, bottom). However, we observed variably low numbers of *SENCR* RNA molecules in the cytoplasm of individual HUVEC (Figure 4A, top and middle). We sometimes observed *SENCR* RNA in the nucleus although this probably reflects either active transcription or unprocessed RNA. The cytoplasmic, low-level expression of *SENCR* RNA contrasts with the higher-level nuclear accumulation of *NEAT1* lncRNA as well as cytoplasmic *PP1B* mRNA (Figure 4A). Biochemical fractionation followed by RT-PCR further documented cytoplasmic localization of *SENCR* in both HUVEC and HCASMC. In contrast, the lncRNAs *NEAT1* and *XIST* show predominantly nuclear accumulation in these cell types (Figure 4B; Figure III in the online-only Data Supplement). We next used 2 distinct probe sets to *SENCR* in HUVEC treated with a control dicer substrate RNA or 2 dicer substrate RNAs targeting different regions of *SENCR* to further demonstrate the specificity of the signal (Figure 4C). Quantitative analysis of coincident hybridization of each probe set demonstrated a likely underestimate of ≈0.8 copies of *SENCR* per cell, a value that was approximately halved on *SENCR* knockdown (Figure 4D). These results establish the cytoplasmic localization of *SENCR* and indicate its relatively weaker level of expression as compared with housekeeping mRNA molecules (*PP1B*) and at least one other lncRNA (*NEAT1*).

SENCR Knockdown Exerts Little Effect on *FLI1* mRNA in Vascular Cells

Many lncRNAs that overlap protein-coding genes in the antisense orientation exert *cis* or *trans* effects on gene expression through the recruitment of chromatin remodeling factors.[52] However, no uniform *cis*-acting effect on *FLI1* or neighboring gene expression was observed on knocking down *SENCR* with multiple dicer substrate RNAs in HCASMC (Figure 5A–5C) or HUVEC (Figure 5D and 5E), consistent with its cytoplasmic localization. There was also little effect of *SENCR* knockdown on the nuclear accumulation of FLI1 protein or steady-state FLI1 protein levels (Figure 6G and 6H). Further,

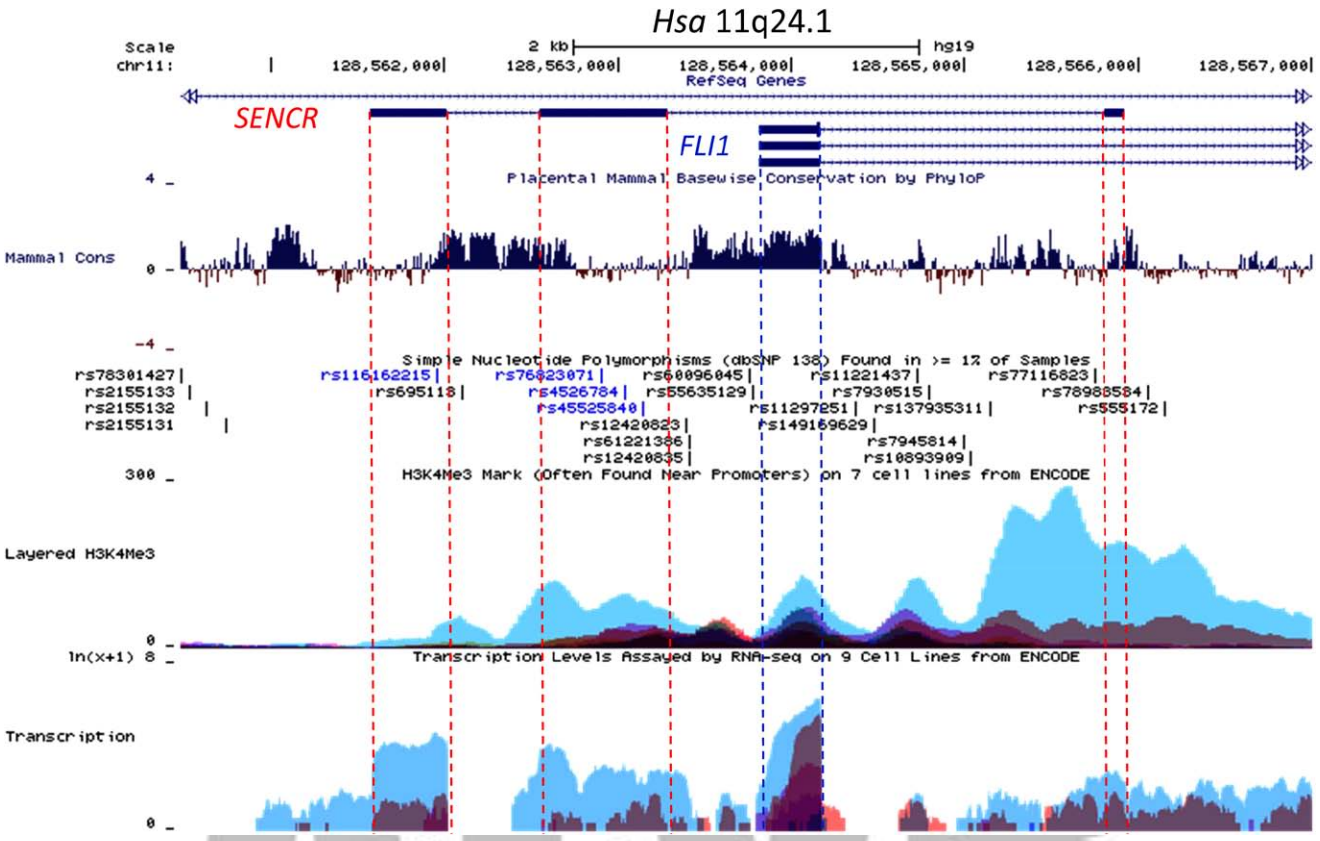


Figure 3. UCSC genome browser track of the human *FLI1-SENCR* sense–antisense gene pair. The *SENCR* gene comprises 3 exons (shown as dark rectangles) and 2 introns. The first exon of *SENCR* initiates on the opposite strand ≈1.5 kb downstream from the first exon of *FLI1*. The dotted vertical lines serve to highlight several features, including (from the top) mammalian conservation (Mammal Cons), reference SNPs (rs numbers), H3K4Me3, and RNA-seq (Transcription) in Tier 1 and Tier 2 cells from Encyclopedia of DNA Elements (ENCODE). Note the lower conservation and selective transcription in human umbilical vein endothelial cell (HUVEC; blue peaks at bottom) for *SENCR* as compared to *FLI1*. See Results for more details.

knockdown of *FLI1* effected no significant change in levels of *SENCR* RNA (Figure 5F). We occasionally observed mild variation in *FLI1* mRNA expression (either up or down) with some dicer substrate RNAs in certain isolates of vascular cells; however, these changes were sporadic and not reproducible when tested by multiple investigators. We therefore conclude that reducing *SENCR* RNA has little to no *cis*-acting effect on local gene expression.

SENCR Knockdown Alters the Normal Contractile Gene Program in HCASMC

Several cytoplasmic lncRNAs effect changes in a cell's transcriptome through post-transcriptional control processes.[53] As an initial step toward understanding the function of *SENCR*, we performed RNA-seq in HCASMC after knockdown of *SENCR* to assess changes in the transcriptome. Most sequencing reads were aligned to the reference genome and scatterplots of replicates showed similar transcript profiles (not shown). Statistical analysis of each set of replicates revealed hundreds of genes that were significantly induced or repressed on *SENCR* knockdown (Figure 6A; Table III in the online-only Data supplement). Strikingly, many SMC contractile genes showed significant reduction in mRNA expression with *SENCR* knockdown (Figure 6B; Table III in the online-only Data supplement). Gene ontology analysis using DAVID revealed biological processes associated with this reduced contractile gene signature (Table IV in the online-only Data supplement). Of note, the key transcriptional switch for SMC contractile gene expression, *MYOCD*,[39] was also reduced with *SENCR* knockdown (Figure 6B), and several dicer substrate RNAs to *SENCR* validated such downregulation in HCASMC (Figure 6D). We also confirmed reduced expression of several of the SMC contractile genes at both the mRNA level (Figure 6E) and the protein level (Figure 6G). Although the SMC contractile program was reduced with *SENCR* knockdown, several genes associated with cell migration were induced (Figure 6C; Table III in the online-only Data Supplement). DAVID analysis supported biological processes linked to cellular locomotion with *SENCR* knockdown (Table V in the online-only Data Supplement). We validated 2 migratory genes (*MDK* and *PTN*) at the mRNA level in HCASMC (Figure 6F) and HUVEC (Figure IV in the online-only Data Supplement). Collectively, these data show that reduced *SENCR* expression compromised the SMC contractile phenotype and promoted a promigratory gene signature.

Attenuated *SENCR* Expression Confers a Hyper-Motile Phenotype in HCASMC

To ascertain whether the increase in promigratory gene expression on knockdown of *SENCR* translates into a

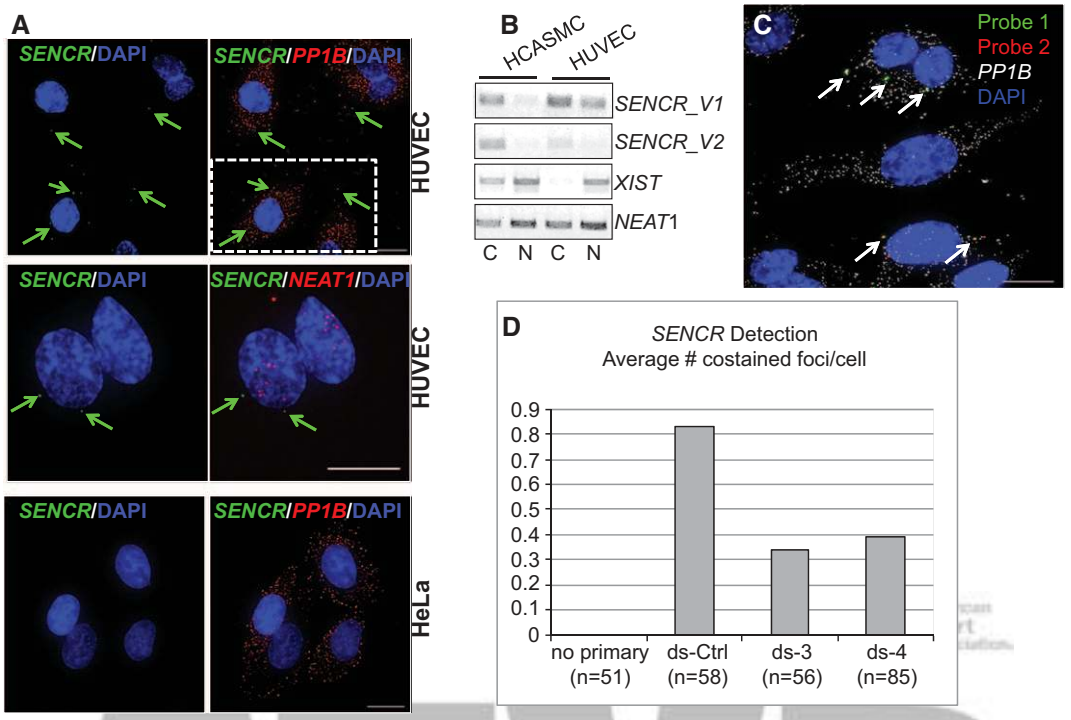


Figure 4. Localization of *SENCR* RNA. **A**, RNA fluorescence in situ hybridization (RNA FISH) analysis of *SENCR* vs cytoplasmic *PP1B* mRNA and nuclear *NEAT1* RNA in indicated cell types. Arrows point to single molecules of *SENCR* RNA in the cytoplasm of human umbilical vein endothelial cell (HUVEC). Scale bars are all 5 μm. The broken white rectangle at **upper right** is shown at higher magnification in Figure III in the online-only Data Supplement. **B**, Reverse transcription polymerase chain reaction analysis of 2 *SENCR* isoforms vs other long noncoding RNA (lncRNA) genes from polyA+RNA isolated from the cytoplasmic (C) or nuclear (N) fractions of indicated cell types. **C**, Application of 2 probe sets to *SENCR* RNA (see Methods in the online-only Data Supplement). Arrows point to coincident localization of 2 fluorescently tagged probe sets (yellow) targeting different regions of the *SENCR* transcript. **D**, Quantitative measures of coincident localization of *SENCR* probes in HUVEC transfected with a dicer substrate control RNA (ds-Ctrl) or 2 dicer substrate RNAs targeting different regions of *SENCR* (ds-3 and ds-4). The *y* axis indicates the average number of costained foci/cell. DAPI indicates 4′,6-diamidino-2-phenylindole.

functional phenotype, we performed 2 independent measures of cell migration. Using a scratch wound assay, we observed hyper-motile HCASMC with *SENCR* knockdown (Figure 7A and 7B). Many of these cells exhibited reorganization of the actin cytoskeleton with formation of lamellipodia, consistent with a migratory cell phenotype (Figure 7A and 7F, arrows). Importantly, the increase in HCASMC migration could be completely rescued on simultaneous knockdown

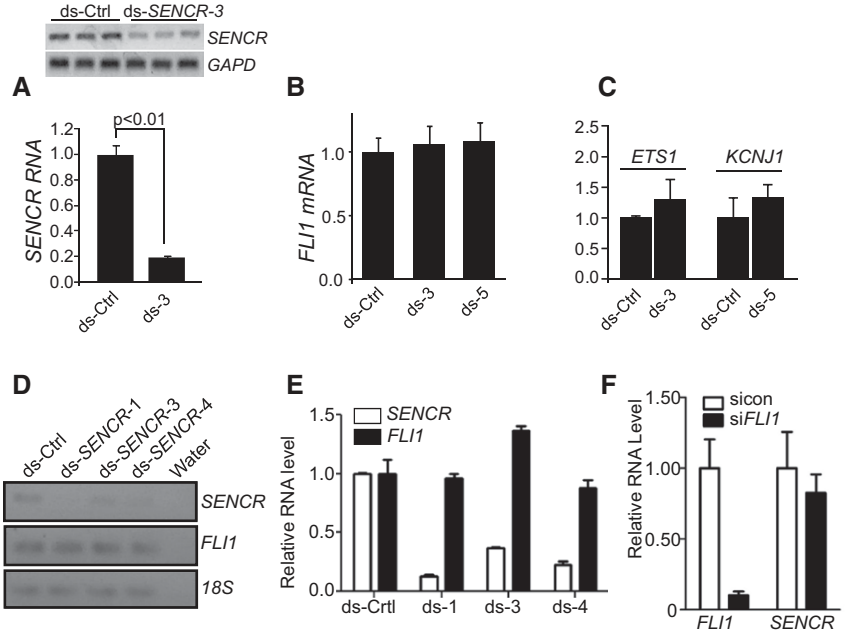
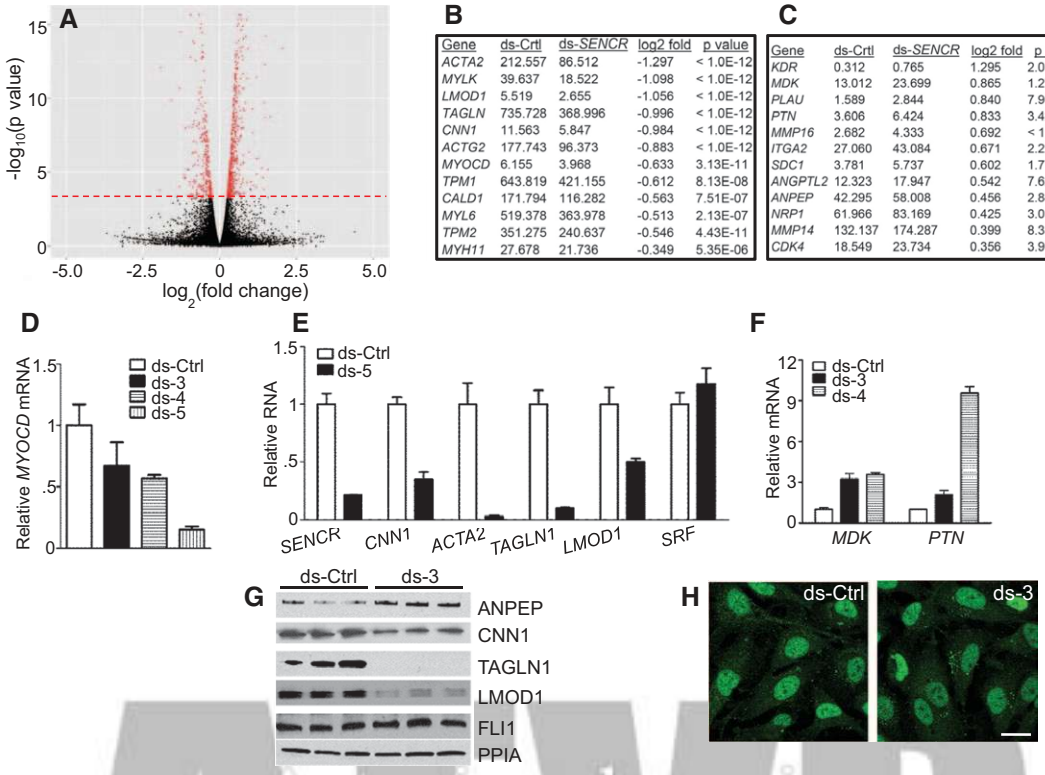


Figure 5. Effect of knocking down *SENCR* on local gene expression. **A**, Dicer substrate control (ds-Ctrl) or dicer substrate *SENCR* RNA was transfected into human coronary artery smooth muscle cell (HCASMC) for 72 hours and then total RNA isolated for conventional (**top**) or quantitative (**bottom**) reverse transcription polymerase chain reaction (RT-PCR). Dicer substrate RNA to various regions of *SENCR* are abbreviated here and below as ds followed by a number (see Table I in the online-only Data Supplement for details). Quantitative RT-PCR of *FLI1* mRNA (**B**) or flanking genes around *FLI1* (**C**) after 3 d transfection with indicated dicer substrate RNAs. **D**, Conventional RT-PCR of *SENCR* and *FLI1* in human umbilical vein endothelial cell (HUVEC) after transfection with indicated dsRNAs. **E**, Quantitative RT-PCR of *SENCR* and *FLI1* after transfection with indicated dsRNAs in HUVEC. **F**, Quantitative RT-PCR of *FLI1* and *SENCR* after knockdown of *FLI1* mRNA in HCASMC. Data are representative of multiple independent experiments performed by independent investigators using several isolates of HCASMC or HUVEC.



Gene	ds-Crtl	ds-*SENCR*	log2 fold	p value
ACTA2	212.557	86.512	-1.297	< 1.0E-12
MYLK	39.637	18.522	-1.098	< 1.0E-12
LMOD1	5.519	2.655	-1.056	< 1.0E-12
TAGLN	735.728	368.996	-0.996	< 1.0E-12
CNN1	11.563	5.847	-0.984	< 1.0E-12
ACTG2	177.743	96.373	-0.883	< 1.0E-12
MYOCD	6.155	3.968	-0.633	3.13E-11
TPM1	643.819	421.155	-0.612	8.13E-08
CALD1	171.794	116.282	-0.563	7.51E-07
MYL6	519.378	363.978	-0.513	2.13E-07
TPM2	351.275	240.637	-0.546	4.43E-11
MYH11	27.678	21.736	-0.349	5.35E-06

Gene	ds-Crtl	ds-*SENCR*	log2 fold	p value
KDR	0.312	0.765	1.295	2.01E-11
MDK	13.012	23.699	0.865	1.27E-14
PLAU	1.589	2.844	0.840	7.96E-08
PTN	3.606	6.424	0.833	3.42E-09
MMP16	2.682	4.333	0.692	< 1.0E-14
ITGA2	27.060	43.084	0.671	2.22E-16
SDC1	3.781	5.737	0.602	1.77E-11
ANGPTL2	12.323	17.947	0.542	7.62E-14
ANPEP	42.295	58.008	0.456	2.84E-09
NRP1	61.966	83.169	0.425	3.01E-06
MMP14	132.137	174.287	0.399	8.35E-06
CDK4	18.549	23.734	0.356	3.96E-05

Figure 6. Effect of *SENCR* knockdown on human coronary artery smooth muscle cell (HCASMC) transcriptome. **A**, Volcano plot depicting changes in gene expression with *SENCR* knockdown. The red dashed line indicates genes (in red) whose changes in expression were statistically significant. Sample smooth muscle cell (SMC) contractile genes (**B**) and promigratory genes (**C**) exhibiting either reduced (**B**) or increased (**C**) expression with dicer substrate (ds)-*SENCR* knockdown. See Table III in the online-only Data Supplement for a complete listing of all genes showing significant up- or downregulation with *SENCR* knockdown. Quantitative reverse transcription polymerase chain reaction (RT-PCR) validation of reduced *MYOCD* mRNA (**D**) and SMC contractile genes (**E**) in HCASMC after knockdown of *SENCR* with various dsRNAs. **F**, Quantitative RT-PCR validation of upregulation of 2 promigratory genes on knockdown of *SENCR* in HCASMC. **G**, Western blot validation of upregulated (ANPEP) and downregulated (SMC contractile) proteins in HCASMC 72 hours after indicated transfection with dsRNA. Similar findings were observed in an independent experiment. **H**, Immunofluorescence microscopy of FLI1 protein in the nucleus of HCASMC after 3 d transfection with indicated dsRNAs. Results are representative of multiple experiments performed by independent investigators. Scale bar, 10 μm for both images. SRF indicates serum response factor.

of either of 2 promigratory genes shown to be induced on knockdown of *SENCR* (Figure 7C; Figure V in the online-only Data Supplement). To further confirm this accentuated cell migration phenotype on knockdown of *SENCR*, we used a modified Boyden chamber assay. Consistent with the scratch wound assay, we noted that HCASMC migration was elevated with *SENCR* knockdown although not as much as that observed with the potent migratory stimulus, PDGF-BB (Figure 7D and 7E). We also observed augmented PDGF-BB–induced cell migration on concomitant knockdown of *SENCR* (Figure VI in the online-only Data Supplement). Taken together, these results strongly support a role for *SENCR* in the regulation of HCASMC differentiation and cellular motility.

Discussion

Contrary to the historical notion of pervasive junk DNA,[1] most of the human genome is transcribed signifying a treasure-trove of previously unrecognized functional DNA sequences. These include tens of millions of regulatory elements as well as the expansive class of lncRNA genes. LncRNA genes already outnumber protein-coding genes and they exhibit diverse functions related to gene expression and splicing; protein translation, activity, and trafficking; as well as the formation of specialized microenvironmental niches.[54,55] Here, we present the first RNA-seq study in a human vascular cell type for the specific discovery of lncRNA genes. We used strict criteria and discovered 31 previously unannotated lncRNAs, 21 of which we validated in human cell lines and human tissues. In addition, we detected a few lncRNAs in dated human plasma, suggesting that these may have potential utility as biomarkers of clinical disease.[56] One of the lncRNA genes, named here as *SENCR*, shows a selective pattern of expression in cells and tissues with highest levels in human vascular SMC and ECs. We discovered that *SENCR* undergoes alternative splicing, consistent with widespread splicing of transcripts across the human genome.[57] *SENCR* overlaps the 5′ end of the *FLI1* transcription factor in the antisense orientation, but does not seem to regulate local gene expression in *cis*. Indeed, our extensive RNA fluorescence in situ hybridization and biochemical fractionation studies clearly indicate *SENCR* to be a cytoplasmic lncRNA supporting an extranuclear function. Using RNA-seq after knockdown of *SENCR*, we observed uniform decreases in expression of SMC contractile–associated genes as well as attenuated expression of the major

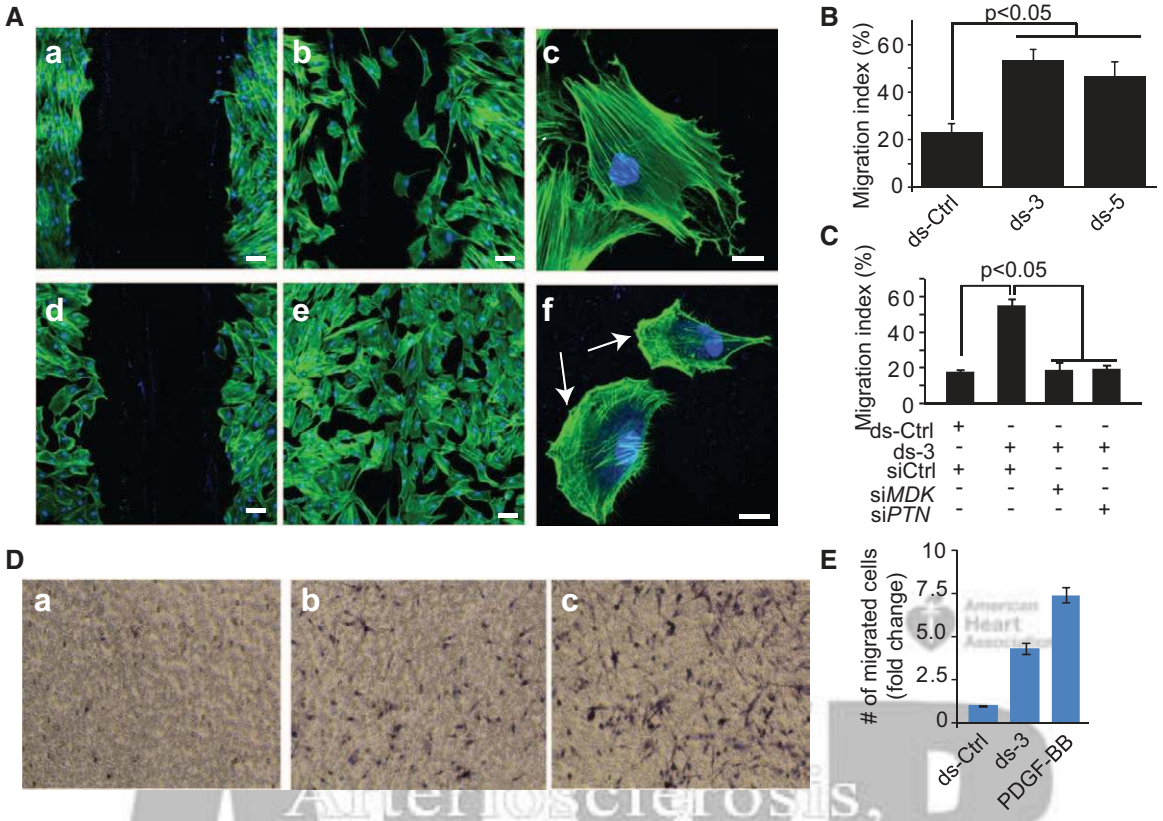


Figure 7. Effect of *SENCR* knockdown in a scratch wound assay of cell migration. **A**, Human coronary artery smooth muscle cells (HCASMCs) were transfected with ds-Ctrl (a–c) or ds-*SENCR*-3 (d–f) for 72 hours after which a scratch wound was created and cell migration assessed in quiescent (a, d) HCASMC or in similar cells stimulated for 12 hours with 10% FBS (b, c, e, and f). Cells were stained with phalloidin (red) and 4′,6-diamidino-2-phenylindole (blue). Bars are 25 μm in a, b, d, and e and 10 μm in c and f. Arrows in f denote lamellipodia. **B**, Quantitative measure of the area of the wound occupied by dsRNA-transfected HCASMC 12 hours after serum stimulation. **C**, Same experiment as in **B** only HCASMC were transfected simultaneously with a control siRNA (siCtrl) or an siRNA to 1 of 2 promigratory genes. Each siRNA reduced level of *MDK* or *PTN* mRNA by >80% (Figure V in the online-only Data Supplement). **D**, Effect of *SENCR* knockdown on HCASMC migration in a Boyden chamber. Cells were transfected with ds-Ctrl (a), ds-3 (b), or 25 ng/mL PDGF-BB (c) for 6 hours and the fold change in number of cells migrating through the porous membrane quantitated (**E**). The data reflect cell counts from 5 independent fields.

transcriptional switch (Myocardin) for the differentiation of vascular SMC.[39] However, knockdown of *SENCR* augments a promigratory gene signature that facilitates heightened SMC migration. Thus, we have uncovered a new vascular cell–enriched lncRNA that seems to function in the maintenance of a normal, nonmotile SMC phenotype.

An analysis of 707 sense–antisense gene pairs annotated in the UCSC genome browser[58] shows diversity in structural orientation, with most lncRNAs representing natural antisense transcripts (47.0%), followed by intronic (18.8%), divergent (16.4%), completely overlapping (7.4%), 5′ overlapping (7.1%), and 3′ overlapping (3.4%) lncRNAs (Table VI in the online-only Data Supplement). Much of what is known about sense–antisense gene pairs relates to natural antisense transcripts and effects on local gene expression through such processes as transcriptional interference, double-stranded RNA-mediated events, or the guidance of chromatin remodeling complexes that repress or enhance protein-coding gene expression in *cis* or *trans*.[33,53,59] *SENCR* falls within the subclass of 5′ overlapping lncRNAs whose exons do not overlap with those of the sense protein-coding (or noncoding) gene. The terminal portion of intron 1 of *SENCR* overlaps a region of high homology, likely representing conserved sequences corresponding to the proximal 5′ promoter of *FLI1*. There is another island of homology within intron 2 of *SENCR*, suggesting *SENCR* could be a precursor for conserved small RNA molecules. Although the second and third exons of *SENCR* overlap the 5′ promoter region of *FLI1,* there is comparatively weak sequence conservation suggesting *SENCR* does not sponge critical DNA-binding transcription factors necessary for *FLI1* mRNA expression (Figure 3). In fact, *SENCR* and *FLI1* seem to be coexpressed in several cells and tissues, including vascular SMC. This is entirely congruent with our inability to show a consistent effect of knocking down either *SENCR* or *FLI1* on the other gene's level of expression. It is interesting to note that there are little, if any, data on expression of FLI1 mRNA and protein in vascular SMC. Further, the functionality of FLI1 in vascular SMC has not been assessed although an EC-specific knockout of *Fli1* showed reduced pericytes and vascular SMC investing the dermal microvasculature.[60] In light of FLI1 expression in vascular SMC as reported here, it will be important to directly assess the role of FLI1 in vascular SMC differentiation and function through conditional gene ablation studies.

We know little as to how sense–antisense gene pairs involving lncRNAs are transcriptionally controlled.

Presumably, divergent (head to head) sense–antisense pairs share a common promoter as has been described for many bidirectionally transcribed protein-coding genes.[61] However, it is completely unclear how other sense–antisense pairs may be transcribed, particularly a lncRNA that is coexpressed with the sense mRNA as shown in this report. Simultaneous expression of *FLI1* and *SENCR* would seem unlikely because of transcriptional collision.[62] How then might *SENCR* and *FLI1* be transcribed? Perhaps there are shared promoter elements that facilitate alternating transcription between *SENCR* and *FLI1*. Consistent with this idea, no *SENCR* promoter activity was detected unless sequences encompassing the *FLI1* 5′ region were included, although the level of activity remained much lower when compared with an EC-restricted promoter (*DLL4*; not shown). Interestingly, a previous report showed undetectable activity of the *FLI1* promoter in cells expressing high levels of *FLI1* mRNA.[63] This could imply there exists a remotely acting enhancer element critical for alternating transcription of *SENCR* and *FLI1*. Another possibility is that *SENCR* and *FLI1* are monoallelically expressed in a mutually exclusive manner.[64] Recently, single-cell RNA-seq analysis demonstrated that as much as 24% of autosomal genes exhibit monoallelic expression thus providing support for this hypothesis.[65] Clearly, a major task for future investigative work will be to elucidate the transcriptional control of *SENCR* and other lncRNAs during vascular cell differentiation or pathological conditions.

Elucidating the function of lncRNAs has been hampered by the absence of any obvious lncRNA sequence code. One approach to begin understanding lncRNA function is to reduce the level of lncRNA expression and then evaluate the transcriptome of a cell type.[66] In this study, we knocked down *SENCR* in HCASMC and found that the contractile phenotype of these cells was attenuated with concomitant increases in several promigratory genes leading to enhanced cell motility. The mechanism for such changes in cell phenotype is unknown at this time; however, because *SENCR* is localized to the cytoplasm it seems unlikely that it acts through direct interaction with DNA or the recruitment of chromatin-modifying complexes to target genes as shown for many nuclear lncRNAs.[53,67] It is more probable that *SENCR* functions in some post-transcriptional capacity to effect the observed changes in gene expression. Because all SMC contractile genes were attenuated with *SENCR* knockdown, a post-transcriptional mechanism would likely involve the targeting of a protein or RNA that is antecedent to the SMC contractile gene program. One possibility would be that *SENCR* sponges a low abundant microRNA that otherwise would function to mute the SMC contractile gene program, similar to what has been shown for linc-MD1 in skeletal muscle.[68] Other potential post-transcriptional mechanisms of action for *SENCR* include stabilization, destabilization, or enhanced ribosomal translation of pivotal RNA transcripts, as proposed for other recently defined lncRNAs.[24,69–71] The results of this study provide a foundation for exploration of these and other possible mechanisms of *SENCR* activity using emerging biochemical tools to analyze lncRNA interactions with other macromolecules in the cytoplasm.[72]

The explosive rise of lncRNAs in human and mouse genomes has profound implications for future research in vascular biology. First, unlike microRNAs, which number ≈1000 and almost universally function through a predictable and well-defined process, lncRNAs number in the tens of thousands and their functions and mechanisms of action will be, arguably, as diverse as those for protein-coding genes. This will necessitate a global effort to define all lncRNAs in the vasculature (especially nonpolyadenylated) under normal and stress-induced conditions and delineate their mode of regulation and function. Second, lncRNAs such as *SENCR* are poorly conserved and lack easily defined sequences that would imply a clear function in blood vessels. The apparent lack of orthologous mouse lncRNA genes such as *SENCR* constrains the extent to which experimental analyses can be done in a rigorous and controlled manner to gain functional insights. However, mouse-specific lncRNAs may have limited translational relevance to the study of human development and disease. Structural similarity between lncRNAs having little sequence homology may, nevertheless, exhibit comparable functions across species.[73,74] In this context, there is a pressing need to gain insight into the structure of lncRNAs to develop lncRNA codes that would facilitate functional classification across species. As a first approximation of the structure of *SENCR*, we used mFold (http://mfold.rna.albany.edu) and RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) and found it to exhibit a stable RNA structure (Figure VII in the online-only Data Supplement) with minimum free energies of −486 and −470 kcal/mol, respectively. Another implication of widespread lncRNA genes will be the need for extreme caution and strategic design in the creation of genetically altered mice, especially when targeting the 5′ end of a gene where inadvertent disruption of other sequences such as lncRNAs is likely to occur. The emergence of precision-guided genome editing (eg, CRISPR/Cas9) will be of great value in this context.[75] Finally, most genetic variation occurs in non–protein-coding sequence space,[76] which is interposed with transcription factor binding sites such as CArG boxes[11] and lncRNAs such as *ANRIL*.[35] Historically, there has been a notable lack of understanding as to how noncoding sequence variations associated with disease perturb function in a cell. Now, with increasing efforts devoted to understanding noncoding sequences, there will be an effort to model human SNPs associated with vascular disease through, for example, CRISPR/Cas9-mediated point mutations in the mouse genome. In this context, it will be important to know whether the sequence variants in exons 2 and 3 of *SENCR* confer differential expression, localization, or function in a disease setting. Altered lncRNA expression of *TIE1-AS1*[46] and *ANRIL*[48] has already been noted in human vascular disease.

In summary, we have developed a rigorous experimental pipeline for the discovery and study of lncRNAs in human vascular cells (Figure I in the online-only Data Supplement). This approach uncovered many previously unrecognized lncRNAs, including the human-specific, vascular cell–selective *SENCR*,

which we show is an alternatively spliced and weakly expressed cytoplasmic 5′ overlapping antisense lncRNA. Loss-of-function studies support the concept of *SENCR* acting as a fine-tuner of the vascular SMC phenotype. Of note, *SENCR* is one of the first 5′ overlapping antisense lncRNAs (as defined here in Table VI in the online-only Data Supplement) to be studied in detail. Future work should aim to elucidate the regulatory control and function of *SENCR* in models of human vascular SMC and EC development as well as disease-associated processes.

Acknowledgments

We gratefully acknowledge the expertise of the University of Rochester Genomics Research Center for performing the RNA-seq experiments and analyzing differences in protein-coding genes as well as generating the volcano plot.

Sources of Funding

This work was supported by grants from the National Institutes of Health (HL62572 and HL091168 to J.M. Miano; MH099452 to D. Zheng and partially by HL111770; and 5P01-CA013106 to D.L. Spector) and the American Heart Association (10SDG3670036 to X. Long and 12POST11950002 to R.D. Bell). J.H. Bergmann was funded by a DAAD postdoctoral fellowship.

Disclosures

None.

References

1. Ohno S. So much "junk" DNA in our genome. *Brookhaven Symp Biol*. 1972;23:366–370.
2. Okazaki Y, Furuno M, Kasukawa T, et al.; FANTOM Consortium; RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*. 2002;420:563–573..
3. Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SP, Gingeras TR. Large-scale transcriptional activity in chromosomes 21 and 22. *Science*. 2002;296:916–919.
4. Carninci P, Kasukawa T, Katayama S, et al.; FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). The transcriptional landscape of the mammalian genome. *Science*. 2005;309:1559–1563..
5. Cheng J, Kapranov P, Drenkow J, et al. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science*. 2005;308:1149–1154.
6. Mercer TR, Gerhardt DJ, Dinger ME, Crawford J, Trapnell C, Jeddeloh JA, Mattick JS, Rinn JL. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat Biotechnol*. 2012;30:99–104.
7. Livyatan I, Harikumar A, Nissim-Rafinia M, Duttagupta R, Gingeras TR, Meshorer E. Non-polyadenylated transcription in embryonic stem cells reveals novel non-coding RNA related to pluripotency and differentiation. *Nucleic Acids Res*. 2013;41:6300–6315.
8. Gerstein MB, Kundaje A, Hariharan M, et al. Architecture of the human regulatory network derived from ENCODE data. *Nature*. 2012;489:91–100.
9. Neph S, Vierstra J, Stergachis AB, et al. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature*. 2012;489:83–90.
10. Zhang X, Odom DT, Koo SH, et al. Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. *Proc Natl Acad Sci USA*. 2005;102:4459–4464.
11. Benson CC, Zhou Q, Long X, Miano JM. Identifying functional single nucleotide polymorphisms in the human CArGome. *Physiol Genomics*. 2011;43:1038–1048.
12. Johnson R, Richter N, Bogu GK, Bhinge A, Teng SW, Choo SH, Andrieux LO, de Benedictis C, Jauch R, Stanton LW. A genome-wide screen for genetic variants that modify the recruitment of REST to its target genes. *PLoS Genet*. 2012;8:e1002624.
13. The Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57–74.
14. Belinky F, Bahir I, Stelzer G, Zimmerman S, Rosen N, Nativ N, Dalah I, Iny Stein T, Rappaport N, Mituyama T, Safran M, Lancet D. Non-redundant compendium of human ncRNA genes in GeneCards. *Bioinformatics*. 2013;29:255–261.
15. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science*. 2001;294:853–858.
16. Falaleeva M, Stamm S. Processing of snoRNAs as a new source of regulatory non-coding RNAs: snoRNA fragments form a new class of functional RNAs. *Bioessays*. 2013;35:46–54.
17. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genet*. 2009;10:94–108..
18. Guttman M, Amit I, Garber M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009;458:223–227.
19. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci USA*. 2009;106:11667–11672.
20. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011;25:1915–1927.
21. Lee JT. Epigenetic regulation by long noncoding RNAs. *Science*. 2012;338:1435–1439.
22. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell*. 2013;154:26–46.
23. Gutschner T, Hämmerle M, Diederichs S. MALAT1—a paradigm for long noncoding RNA function in cancer. *J Mol Med (Berl)*. 2013;91:791–801.
24. Kretz M, Siprashvili Z, Chu C, et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature*. 2013;493:231–235.
25. St Laurent G, Shtokalo D, Dong B, Tackett MR, Fan X, Lazorthes S, Nicolas E, Sang N, Triche TJ, McCaffrey TA, Xiao W, Kapranov P. VlincRNAs controlled by retroviral elements are a hallmark of pluripotency and cancer. *Genome Biol*. 2013;14:R73.
26. De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol*. 2010;8:e1000384.
27. Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, Guigo R, Shiekhattar R. Long noncoding RNAs with enhancer-like function in human cells. *Cell*. 2010;143:46–58.
28. Tahira AC, Kubrusly MS, Faria MF, Dazzani B, Fonseca RS, Maracaja-Coutinho V, Verjovski-Almeida S, Machado MC, Reis EM. Long noncoding intronic RNAs are differentially expressed in primary and metastatic pancreatic cancer. *Mol Cancer*. 2011;10:141.
29. St Laurent G, Shtokalo D, Tackett MR, Yang Z, Eremina T, Wahlestedt C, Urcuqui-Inchima S, Seilheimer B, McCaffrey TA, Kapranov P. Intronic RNAs constitute the major fraction of the non-coding RNA in mammalian cells. *BMC Genomics*. 2012;13:504.
30. Chen J, Sun M, Kent WJ, Huang X, Xie H, Wang W, Zhou G, Shi RZ, Rowley JD. Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res*. 2004;32:4812–4820.
31. Katayama S, Tomaru Y, Kasukawa T, et al.; RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium. Antisense transcription in the mammalian transcriptome. *Science*. 2005;309:1564–1566..
32. Derrien T, Johnson R, Bussotti G, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res*. 2012;22:1775–1789.
33. Magistri M, Faghihi MA, St Laurent G 3rd, Wahlestedt C. Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. *Trends Genet*. 2012;28:389–396.
34. Lee JT, Bartolomei MS. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell*. 2013;152:1308–1323.
35. Pasmant E, Sabbagh A, Vidaud M, Bièche I. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J*. 2011;25:444–448.
36. Volders PJ, Helsens K, Wang X, Menten B, Martens L, Gevaert K, Vandesompele J, Mestdagh P. LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res*. 2013;41(Database issue):D246–D251..
37. Norman C, Runswick M, Pollock R, Treisman R. Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. *Cell*. 1988;55:989–1003.

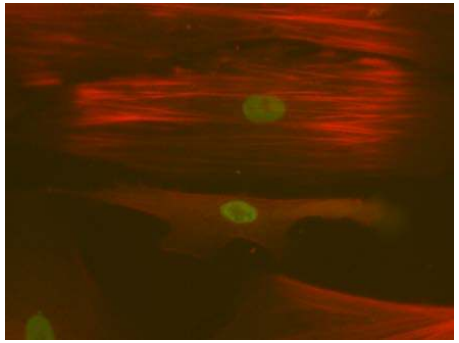
38. Wang D, Chang PS, Wang Z, Sutherland L, Richardson JA, Small E, Krieg PA, Olson EN. Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell*. 2001;105:851–862.
39. Chen J, Kitchen CM, Streb JW, Miano JM. Myocardin: a component of a molecular switch for smooth muscle differentiation. *J Mol Cell Cardiol*. 2002;34:1345–1356.
40. Kume T. Foxc2 transcription factor: a newly described regulator of angiogenesis. *Trends Cardiovasc Med*. 2008;18:224–228.
41. Lammerts van Bueren K, Black BL. Regulation of endothelial and hematopoietic development by the ETS transcription factor Etv2. *Curr Opin Hematol*. 2012;19:199–205.
42. De Val S, Chi NC, Meadows SM, Minovitsky S, Anderson JP, Harris IS, Ehlers ML, Agarwal P, Visel A, Xu SM, Pennacchio LA, Dubchak I, Krieg PA, Stainier DY, Black BL. Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. *Cell*. 2008;135:1053–1064.
43. Small EM, Olson EN. Pervasive roles of microRNAs in cardiovascular biology. *Nature*. 2011;469:336–342.
44. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem*. 2012;81:145–166.
45. Han DK, Khaing ZZ, Pollock RA, Haudenschild CC, Liau G. H19, a marker of developmental transition, is reexpressed in human atherosclerotic plaques and is regulated by the insulin family of growth factors in cultured rabbit smooth muscle cells. *J Clin Invest*. 1996;97:1276–1285.
46. Li K, Blum Y, Verma A, et al. A noncoding antisense RNA in tie-1 locus regulates tie-1 function in vivo. *Blood*. 2010;115:133–139.
47. Congrains A, Kamide K, Katsuya T, Yasuda O, Oguro R, Yamamoto K, Ohishi M, Rakugi H. CVD-associated non-coding RNA, ANRIL, modulates expression of atherogenic pathways in VSMC. *Biochem Biophys Res Commun*. 2012;419:612–616.
48. Motterle A, Pu X, Wood H, Xiao Q, Gor S, Ng FL, Chan K, Cross F, Shohreh B, Poston RN, Tucker AT, Caulfield MJ, Ye S. Functional analyses of coronary artery disease associated variation on chromosome 9p21 in vascular smooth muscle cells. *Hum Mol Genet*. 2012;21:4021–4029.
49. Leung A, Trac C, Jin W, Lanting L, Akbany A, Sætrom P, Schones DE, Natarajan R. Novel long noncoding RNAs are regulated by angiotensin II in vascular smooth muscle cells. *Circ Res*. 2013;113:266–278.
50. Liu F, Walmsley M, Rodaway A, Patient R. Fli1 acts at the top of the transcriptional network driving blood and endothelial development. *Curr Biol*. 2008;18:1234–1240.
51. Battich N, Stoeger T, Pelkmans L. Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat Methods*. 2013;10:1127–1133.
52. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell*. 2009;136:629–641.
53. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet*. 2014;15:7–21..
54. Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. *Cell*. 2013;152:1298–1307.
55. Kung JT, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics*. 2013;193:651–669.
56. Arita T, Ichikawa D, Konishi H, Komatsu S, Shiozaki A, Shoda K, Kawaguchi T, Hirajima S, Nagata H, Kubota T, Fujiwara H, Okamoto K, Otsuji E. Circulating long non-coding RNAs in plasma of patients with gastric cancer. *Anticancer Res*. 2013;33:3185–3193.
57. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet.* 2008;40:1413–1415.
58. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res*. 2002;12:996–1006.
59. Lapidot M, Pilpel Y. Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO Rep*. 2006;7:1216–1222.
60. Asano Y, Stawski L, Hant F, Highland K, Silver R, Szalai G, Watson DK, Trojanowska M. Endothelial Fli1 deficiency impairs vascular homeostasis: a role in scleroderma vasculopathy. *Am J Pathol*. 2010;176:1983–1998.
61. Wakano C, Byun JS, Di LJ, Gardner K. The dual lives of bidirectional promoters. *Biochim Biophys Acta*. 2012;1819:688–693..
62. Hobson DJ, Wei W, Steinmetz LM, Svejstrup JQ. RNA polymerase II collision interrupts convergent transcription. *Mol Cell.* 2012;48:365–374.
63. Barbeau B, Bergeron D, Beaulieu M, Nadjem Z, Rassart E. Characterization of the human and mouse Fli-1 promoter regions. *Biochim Biophys Acta*. 1996;1307:220–232.
64. Raslova H, Komura E, Le Couédic JP, Larbret F, Debili N, Feunteun J, Danos O, Albagli O, Vainchenker W, Favier R. FLI1 monoallelic expression combined with its hemizygous loss underlies Paris-Trousseau/Jacobsen thrombopenia. *J Clin Invest*. 2004;114:77–84.
65. Deng Q, Ramsköld D, Reinius B, Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science*. 2014;343:193–196.
66. Panzitt K, Tschernatsch MM, Guelly C, Moustafa T, Stradner M, Strohmaier HM, Buck CR, Denk H, Schroeder R, Trauner M, Zatloukal K. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology*. 2007;132:330–342.
67. Yang L, Froberg JE, Lee JT. Long noncoding RNAs: fresh perspectives into the RNA world. *Trends Biochem Sci*. 2014;39:35–43.
68. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, Tramontano A, Bozzoni I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*. 2011;147:358–369.
69. Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, Finch CE, St Laurent G 3rd, Kenny PJ, Wahlestedt C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med.* 2008;14:723–730..
70. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature*. 2011;470:284–288.
71. Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, Pesce E, Ferrer I, Collavin L, Santoro C, Forrest AR, Carninci P, Biffo S, Stupka E, Gustincich S. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature*. 2012;491:454–457.
72. Zhu J, Fu H, Wu Y, Zheng X. Function of lncRNAs and approaches to lncRNA-protein interactions. *Sci China Life Sci*. 2013;56:876–885.
73. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*. 2011;147:1537–1550.
74. Torarinsson E, Sawera M, Havgaard JH, Fredholm M, Gorodkin J. Thousands of corresponding human and mouse genomic regions unalignable in primary sequence contain common RNA structure. *Genome Res*. 2006;16:885–889.
75. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;153:910–918.
76. Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. 2012;337:1190–1195.

Significance

For the first time, RNA-seq has been performed in human coronary artery smooth muscle cell for the discovery of long noncoding RNA genes. We report the gene structure, expression, splicing, and spatial localization of a new vascular cell–selective long noncoding RNA we call *SENCR*. Although *SENCR* has no apparent *cis* effect on gene expression, there is a compromise in the smooth muscle cell contractile gene program on its knockdown with elevations in many promigratory genes. Accordingly, these cells exhibit a hyper-motile phenotype, which can be reversed by knocking down 2 promigratory genes that are induced with *SENCR* knockdown. These results report the first novel long noncoding RNA gene selectively expressed in human vascular cells and provide a framework for further study of long noncoding RNA genes during vascular cell development and in disease processes.

Supplemental Figure I

HCASMC (CNN1/SRF)



RNA-seq
oligo-dT primed, single end, 20 million reads

Bioinformatic Analyses
TopHat, Scripture, Cufflinks, Cuffdiff

31 lncRNAs
Non-annotated, non-protein coding,
≥ 2 exons, ≥ 200 bp, ≥ 0.7 FPKM

Validate Expression/Strand
(q)RT-PCR, RACE, Sequencing

Cell/Tissue Profiling/Splicing
(q)RT-PCR, exon-specific hybridize
PCR products, sequencing

Assess *cis* Effects
SENCR knockdown, (q)RT-PCR

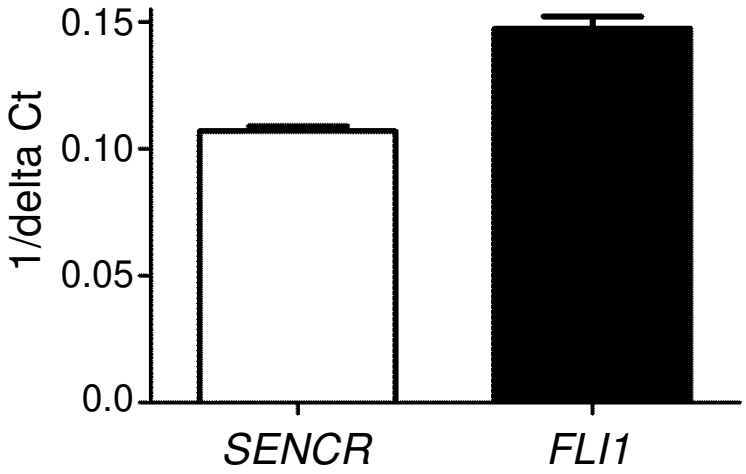
Cellular Localization
Cell Fractionation, RNA-FISH

Bioinformatics
miR sites, 2º structure

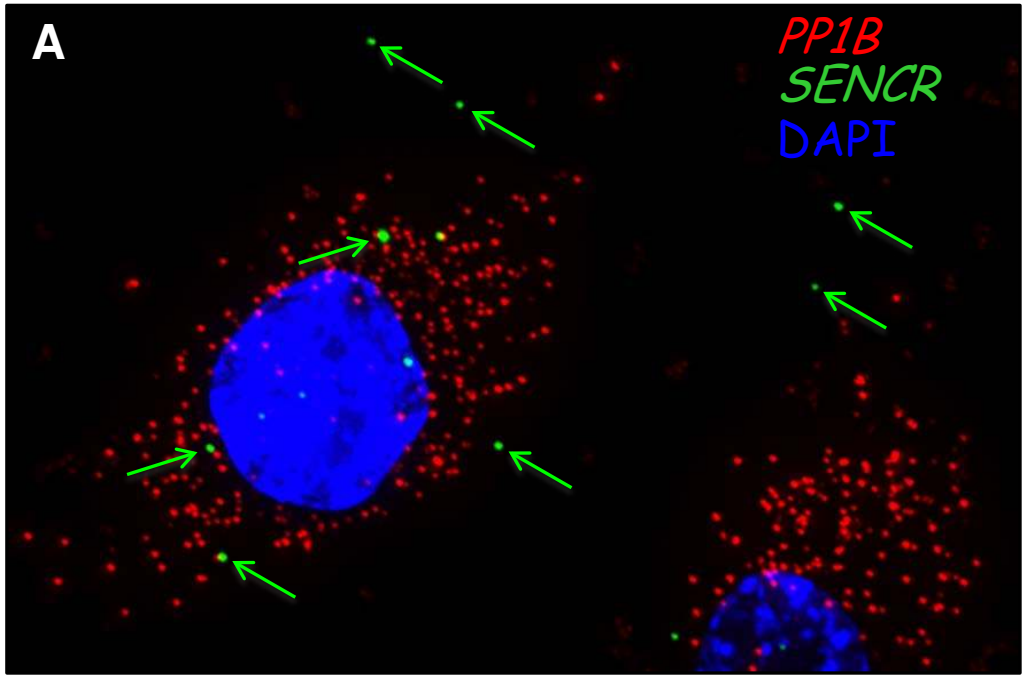
Functional Studies
SENCR knockdown > RNA-seq
- validate (RT-PCR/Western)
- GO analysis (DAVID)
- Migration assay (Rescue)

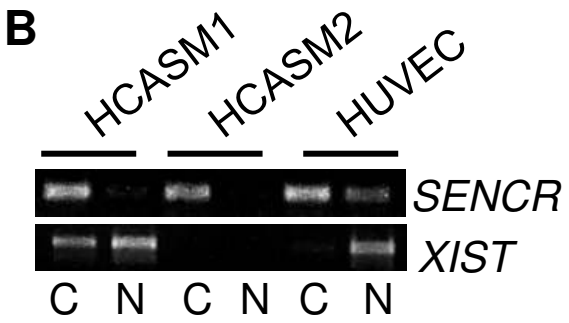
Promoter activity (Luciferase)

Supplemental Figure I. Summary of experimental workflow. We developed this workflow for the study of *SENCR* and other unannotated or uncharacterized lncRNAs. Primary cultures of HCASMC were chosen that express contractile proteins (such as CNN1 in red) as well as the SRF transcription factor (shown in green). See Materials and Methods and Results for further details.

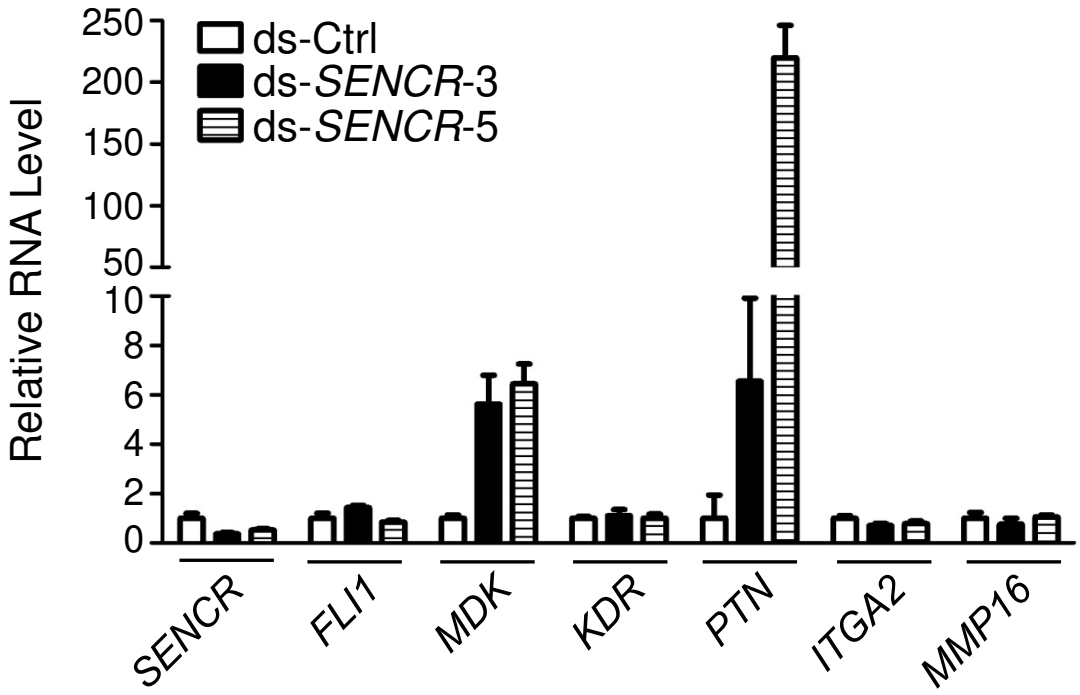


Supplemental Figure II. Relative level of *SENCR* versus *FLI1* in HCASMC. One divided by the delta Ct value for *SENCR* (n=9) and *FLI1* (n=9) RNA expression in HCASMC.

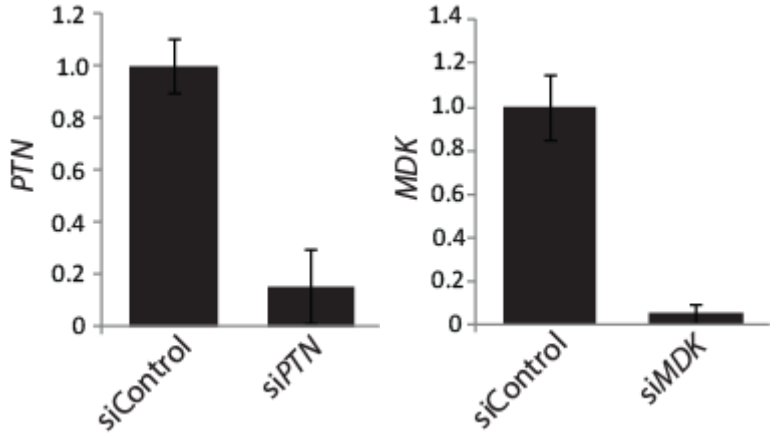




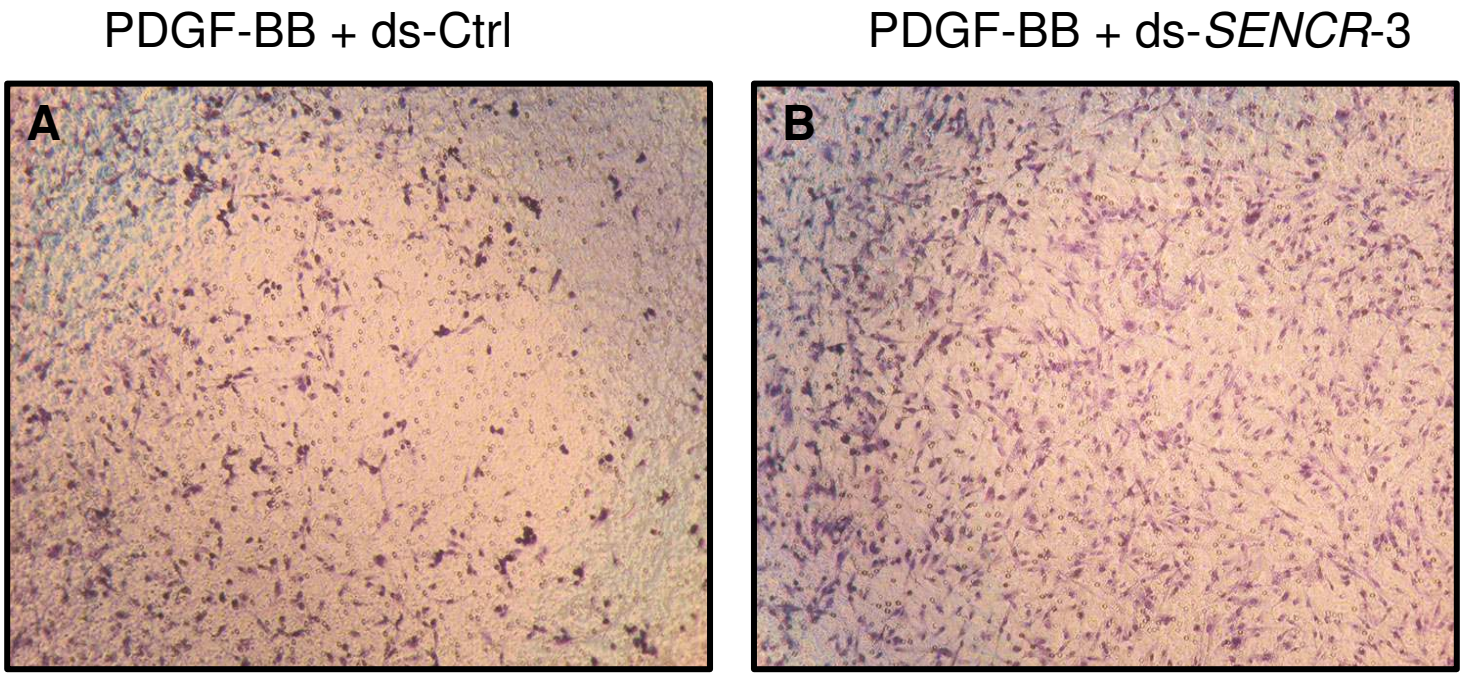
Supplemental Figure III. Localization of *SENCR*. (A) Higher magnification image of boxed region in Figure 4A. Arrows indicate *SENCR* transcripts localized to the cytoplasm of two HUVEC. Omission of labeled probes revealed no background fluorescence (not shown). (B) *SENCR* versus *XIST* RNA localization in cytoplasmic (C) or nuclear (N) fractions of the indicated cell types.



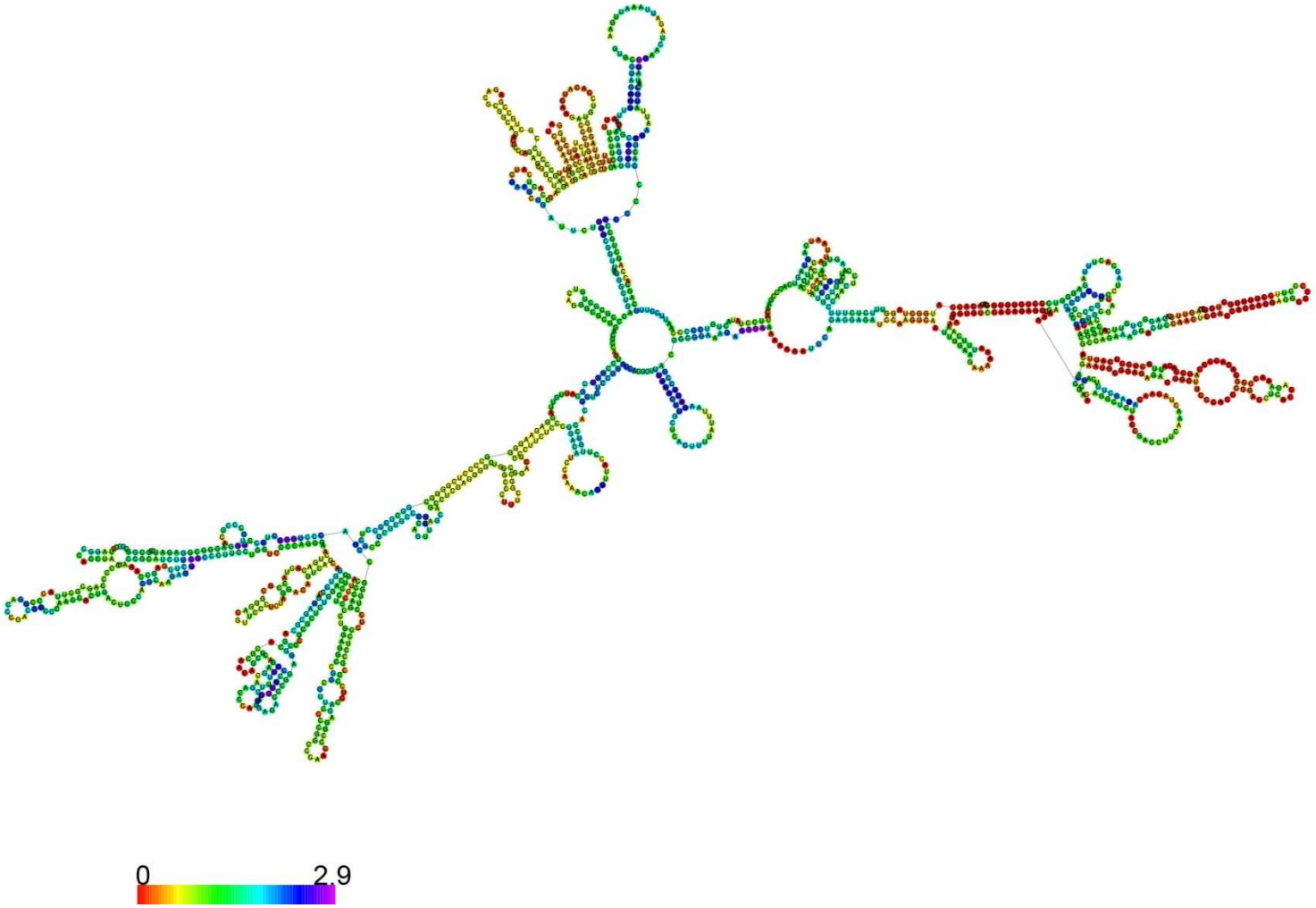
Supplemental Figure IV. Effect of *SENCR* knockdown on expression of pro-migratory genes in HUVEC. Quantitative RT-PCR of indicated genes following transfection with either control dicer substrate RNA or either of two dicer substrate RNAs that target different regions of *SENCR*. Note obvious reductions in *SENCR* upon its targeted knockdown with minimal effects on *FLI1*, but associated induction of *PTN* and *MDK*.



Supplemental Figure V. siRNA knockdown of *PTN* and *MDK*. Quantitative RT-PCR analysis of *PTN* (**left**) and *MDK* (**right**) mRNA levels after siRNA knockdown. The y-axis represents the normalized levels of each transcript with si-Controls set to 1.



Supplemental Figure VI. Effect of combined PDGF-BB treatment and *SENCR* knockdown on HCASMC migration. Boyden chamber assay with HCASMC transfected with either control (**A**) or *SENCR* (**B**) dicer substrate RNA followed by 6 hr stimulation with 25 ng/ml PDGF-BB. Note accentuated HCASMC migration with combined PDGF-BB/*SENCR* knockdown.



Supplemental Figure VII. Predicted secondary structure of *SENCR_V1*. This secondary structure was generated with the program RNAFold. See Discussion for more details.

Materials and Methods

Cells and Tissues – Several independent isolates of primary HCASMC and human umbilical vein endothelial cells (HUVEC) were maintained in growth medium supplied by the manufacturer (Invitrogen). HUVEC were obtained from the Cell Culture Core in the Aab CVRI and plated onto gelatin-coated plates/chambers. HeLa, HEK293, SKLMS (a human leiomyosarcoma cell line of uterine SMC origin), LnCAP, and MCF7 cells were grown in medium as specified by the manufacturer (ATCC). Human tissue RNA samples were obtained from a commercial source (Zyagen). Dated human plasma was obtained through the University of Rochester Medical Center Blood Bank.

RNA-Sequencing Analysis – Total RNA was isolated from HCASMC using RNeasy extraction kit (Qiagen) under normal growth conditions or where *SENCR* was knocked down for 3 days with 25 nM of either a dicer substrate RNA to exon 2 (ds-*SENCR*-5, Table I in the online only Data supplement) or a control dicer substrate RNA. Following bioanalyzer quality control confirmation, RNA-seq was performed on the polyadenylated fraction using Illumina Genome Analyzer IIx platform at the University of Rochester Medical Center Genomics Research Center (http://www.urmc.rochester.edu/fgc/). Single-end sequencing was done at a depth of 20 million reads per replicate (n=3). Pre-processing of raw sequence reads included demultiplexing with CASAVA 1.8.2, transcript trimming of contaminating sequences with Sequence Cleaner (http://sourceforge.net/projects/seqclean/), removal of vector sequences with UniVec database (http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html), and FASTQ quality trimming using the FASTX Toolkit (http://cancan.cshl.edu/labmembers/gordon/fastx_toolkit/index.html). SHRiMP2.2.3 was used to align sequence reads to annotated transcripts on the UCSC Reference Genome (build GRCh37/hg19). Quantitative analysis, including the statistical analysis of differentially expressed genes was done with Cufflinks 2.0.2 and Cuffdiff2 (http://cufflinks.cbcb.umd.edu). For the *SENCR* knockdown RNA-seq experiment, the Benjamini-Hochberg method was applied for multiple test correction (FDR < 0.05). Data output files such as Volcano plots and scatterplots were generated with cummerbund (http://compbio.mit.edu/cummeRbund/). Gene ontology (GO) analysis was done using DAVID [1]. All RNA-seq data were deposited into NCBI's Gene Expression Omnibus (GSE51878).

Bioinformatics Methods for Identifying Novel LncRNAs – RNA-seq reads were aligned to the human genome (hg19) using TopHat 1.4 [2]. In this analysis, two iterations of TopHat alignment were performed in order to maximize the chance of identification of exon-exon junctions. The alignment data were used to define novel lncRNAs following the method described for lncRNA identification [3]. The aligned data for each sample were used independently by two complementary programs, Scripture [4] and Cufflinks [5], for assembling transcripts independent of gene annotation. We determined the threshold for the read coverage of each transcript across all samples by optimizing the sensitivity and specificity for identifying full length versus partial length transcripts of protein coding genes in RefSeq [3]. In the end, we kept assembled transcripts present in both Scripture and Cufflinks outputs, and with ≥2 exons, ≥200 bp and ≥ 2.7 read coverage as reads below this threshold were deemed unreliable in predicting exon structure. Next, we eliminated all transcripts that had an exon overlapping in the same strand with known transcripts from available databases. We then computed the coding potential of all remaining putative novel transcripts using PhyloCSF [6] and removed any transcripts containing an open reading frame with PhyloCSF score ≥ 100 across any of three reading frames. We further removed transcripts that were homologous to known protein coding domains in the Pfam database (release 26; both PfamA and PfamB) [7] using the program HMMER-3 (e-value = 10) [8]. Lastly, we computed expression values (in FPKM, fragments per kilobase of exon per million fragments mapped) of all remaining transcripts, together with all

coding and known non-coding genes. The final list of lncRNAs (Table II in the online only Data supplement) comprises transcripts with FPKM >0.7.

Dicer-Substrate RNA Knockdown – Several dicer-substrate RNA (dsRNA) molecules to different exons of *SENCR* or control dsRNA (ds-Ctrl) were synthesized (Integrated DNA Technologies) and pre-tested in HCASMC and HUVEC (see Table I in the online only Data supplement for list of all DNA molecules used in this study). The ds-Ctrl does not target known transcribed sequences in human, mouse, or rat genomes and thus serves as a negative control for dsRNA transfections. Briefly, cells were Lipofectamine-transfected with each dsRNA (20-30 nM) for three days and then total RNA or protein was isolated for further analysis. Results were confirmed using at least two independent dsRNA constructs to *SENCR*, in up to five independent isolates of HCASMC and HUVEC, often times by multiple investigators.

Gene Expression Assays – Total RNA was isolated using the RNeasy kit (Qiagen). RNA integrity was assessed by spectrophotometery (NanoDrop, Thermo Scientific) and agarose gel electrophoresis. cDNA was synthesized from 1μg of total RNA using iScript (Bio-Rad) plus random decamers and/or an oligodT primer. RT-PCR was performed using Platinum PCR Supermix (Invitrogen) with a MyCycler thermocycler (BioRad) and PCR products were resolved in a 1% agarose gel. Some gels shown were adjusted uniformly in Photoshop using the "invert" function. For lncRNA validations, we included a no RT step that revealed little to no products indicating authentic polyA+ RNAs were amplified as opposed to contaminating genomic DNA (data not shown). Quantitative RT-PCR was performed using IQ SYBR Green Supermix with a MyiQ single color real-time PCR detection system (BioRad). Experiments shown are representative of multiple independent experiments using different lots of HCASMC and HUVEC, performed by separate investigators to ensure quality control and accurate interpretation of observed changes in gene expression.

Western Blotting – Total protein was isolated from HCASMC following ds-*SENCR* or ds-Ctrl knockdown using RIPA buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 0.1% SDS, 1.0% NP-40, 0.5% sodium deoxycholate and Roche protease inhibitor cocktail) and resolved in acrylamide gels for Western blotting as previously done [9]. Antibodies were LMOD1 (ProteinTech, 1:2000), TAGLN (Abcam, 1:4000), CNN1 (DAKO, 1:2000), ANPEP (R&D, 1:1000), FLI1 (Santa Cruz, 1:250) and PPIA (Santa Cruz, 1:2000), used as an internal control.

RNA Fluorescence In Situ Hybridization (FISH) – RNA FISH with single-molecule sensitivity was performed using QuantiGene® (QG) ViewRNA ISH Cell Assay reagents (Affymetrix) based on branched DNA technology [10]. Custom probe oligonucleotide pair pools specific for *SENCR* long and short isoforms were designed and synthesized by Affymetrix as "Type 6" (50 pairs targeting all 3 exons; product ID, VA6-14704) and "Type 4" (19 pairs excluding exon 2; product ID, VA4-14958), respectively. A probe pair pool specific for human *PP1B* housekeeping mRNA (VA1-10148, "Type 1") and the lncRNA, *NEAT1* (VA1-12621, "Type 1") were used as cytoplasmic and nuclear controls, respectively, to assist in interpreting spatial localization of *SENCR* RNA. HUVEC (± *SENCR* knockdown) were grown on acid-washed #1.5 glass cover slips (Thermo Scientific) to 70%-80% confluence, washed with PBS, and fixed for 30 min in fresh 0.45μm-filtered 4% paraformaldehyde (Electron Microscopy Sciences) dissolved in Dulbecco's PBS without $CaCl_2$ and $MgCl_2$ (Invitrogen). RNA-FISH was performed with minor modifications from the manufacturer's protocol as follows: After permeabilization using QG Detergent Solution, cells were treated with 0.5% Triton X-100/PBS for 5 min at room temperature. Partial protease digestion was carried out with a 1:6,000 dilution of QG Protease K for 10 min at room temperature. Coverslips were incubated with primary probe pair sets (3-color multiplexing) or QG Probe Set Diluent as negative control at 40°C for 3 hr. Pre-amplifiers were incubated for an extended period of 1 hr. Between probe set incubations, cells were washed 4 times each in QG Wash Buffer for a total of 10 min. After counter-staining with DAPI, coverslips were mounted in home-made anti-fade mounting medium (www.spectorlab.cshl.edu/protocols) and sealed with nail polish.

Image Acquisition and Analysis – Cells were imaged on a DeltaVision Core system (Applied Precision) based on an inverted IX-71 microscope stand (Olympus) equipped with a 60x U-PlanApo 1.40 NA oil immersion lens (Olympus). Images were captured using a CoolSNAP HQ CCD camera (Photometric) as 10μm image stacks with a z-spacing of 0.2μm at a 1x1 binning. Stage, shutter and exposure were controlled through SoftWorx (Applied Precision). Image deconvolution was performed using SoftWorx. Parameters for acquisition and post-acquisition processing were identical for all coverslips. Analysis was done on individual image stacks in 3D space by counting the number of *SENCR* hybridization signals divided by the number of cells in each field of view (≥50 cells in ≥10 randomly chosen fields per experiment). In some experiments, we employed two fluorescently-tagged probe sets (above) in the absence or presence of *SENCR* knockdown to further confirm spatial localization. Only signals that showed overlap of QG "Type 4" and "Type 6" probe sets were considered, thus minimizing potential false-positive signal counts when using single color analysis.

Luciferase Assay – The putative promoter of *SENCR* was defined through 5' RACE (Ambion). Several constructs of varying 5' and 3' length were PCR amplified from HCASMC genomic DNA, cloned into the pGL3 Basic Vector (Promega), and sequence confirmed (URMC Genomics Research Center). HUVEC were plated in 12-well dishes and grown to ~60%-70% confluency and transfected with various *SENCR* promoter constructs or a *DLL4* reporter gene as a positive control. Lipofectamine was used in transfections and the normalized average luciferase activity calculated for each reporter plasmid.

Migration Assays – HCASMC were plated onto coverslips and transfected with either ds-Ctrl or dsRNAs targeting non-overlapping sequences in *SENCR*. Sixty hr after transfection, cells were "scratched" with a sterile P200 pipette tip and the culture medium immediately changed to DMEM containing 10% FBS. 12 hr after scratch wounding, the cells were fixed and stained with Alexa Fluor® 660 Phalloidin and DAPI (Invitrogen) according to manufacturer's instructions. The cells were then imaged by confocal microscopy (Olympus FV1000) and the migratory index (percentage of cells that migrated into the time 0 wound area) was calculated using NIH Image J software. An independent assay for migration was done using a modified Boyden chamber (Corning). Briefly, HCASMC were transfected for three days with either ds-Ctrl or ds-*SENCR* (25 nM) and then seeded into a 24-well Boyden chamber plate. Cells were then serum-deprived overnight and subsequently treated either with PDGF-BB (25 ng/ml) or vehicle for 6 hr. Cells were then fixed, stained with hematoxylin, and imaged with an inverted phase contrast microscope. Migration assays are representative of multiple experiments performed independently by two authors (RDB and XL).

Statistical Analysis – Student's t-test or one way ANOVA followed by Tukey's post-hoc test were used to determine statistical significance of the means (± standard deviation) and graphs were plotted (Graph-Pad Prism 5.0). Statistical significance was assumed at $p < 0.05$.

1. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. David: Database for annotation, visualization, and integrated discovery. *Genome Biology*. 2003;4:3
2. Trapnell C, Pachter L, Salzberg SL. Tophat: Discovering splice junctions with rna-seq. *Bioinformatics*. 2009;25:1105-1111
3. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding rnas reveals global properties and specific subclasses. *Genes Dev*. 2011;25:1915-1927
4. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, Fan L, Koziol MJ, Gnirke A, Nusbaum C, Rinn JL, Lander ES, Regev A. Ab initio reconstruction of cell

type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincrnas. *Nature Biotechnology*. 2010;28:503-510
5. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by rna-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*. 2010;28:511-515
6. Lin MF, Jungreis I, Kellis M. Phylocsf: A comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics*. 2011;27:i275-282
7. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A. The pfam protein families database. *Nucleic acids research*. 2008;36:D281-288
8. Finn RD, Clements J, Eddy SR. Hmmer web server: Interactive sequence similarity searching. *Nucleic acids research*. 2011;39:W29-37
9. Nanda V, Miano JM. Leiomodin 1: A new serum response factor-dependent target gene expressed preferentially in differentiated smooth muscle cells. *The Journal of biological chemistry*. 2012;287:2459-2467
10. Battich N, Stoeger T, Pelkmans L. Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nature methods*. 2013;10:1127-1133