Published in final edited form as:
Arterioscler Thromb Vasc Biol. 2014 June ; 34(6): 1249–1259. doi:10.1161/ATVBAHA.114.303240.

Identification and Initial Functional Characterization of a Human Vascular Cell Enriched Long Non-Coding RNA

Robert D. Bell[1,*], **Xiaochun Long**[1,*], **Mingyan Lin**[2,*], **Jan H. Bergmann**[3], **Vivek Nanda**[1], **Sarah L. Cowan**[1], **Qian Zhou**[1], **Yu Han**[1], **David L. Spector**[3], **Deyou Zheng**[2,4], and **Joseph M. Miano**[1]

[1]Aab Cardiovascular Research Institute, University of Rochester School of Medicine and Dentistry, 601 Elmwood Avenue, Rochester, NY 14642

[2]Department of Genetics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461

[3]Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, NY 11724

[4]Departments of Neurology and Neuroscience, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461

Abstract

Objective—Long non-coding 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 we call *SENCR* (Smooth muscle and Endothelial cell enriched migration/differentiation-associated long Non-Coding RNA). Strand-specific RT-PCR and rapid amplification of cDNA ends indicate that *SENCR* is transcribed antisense from the 5' end of the *FLI1* gene and exists as two 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 following *SENCR* knockdown disclose decreased expression of Myocardin and numerous smooth muscle contractile genes, while a number of pro-migratory genes are increased. RT-PCR and Western blotting experiments validate several differentially expressed genes following *SENCR* knockdown. Loss-of-function studies in scratch wound and Boyden chamber assays support *SENCR* as an inhibitor of smooth muscle cell migration.

To whom correspondence should be addressed: Joseph M. Miano, Aab Cardiovascular Research Institute,University of Rochester School of Medicine & Dentistry, 601 Elmwood Avenue, Rochester, NY, USA, Tel.: (585) 276-9789; Fax: (585) 276-9830; j.m.miano@rochester.eduor Deyou Zheng, Department of Neurology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY, USA, Tel.: (718) 839-7215; deyou.zheng@einstein.yu.edu.
*Equal authorship

Disclosures
None

Conclusion—*SENCR* is a new vascular cell-enriched, cytoplasmic lncRNA that appears to stabilize the smooth muscle cell contractile phenotype.

Keywords

long non-coding RNA; RNA-seq; smooth muscle cell; endothelial cell; migration

Introduction

Although the human genome is thirty times 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 over the last ten 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 (ENCODE) Consortium and many other labs 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%–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 non-coding RNA (ncRNA) in a cell functions to orchestrate basic translation (transfer and ribosomal RNA); however, two 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 utilize Argonaute proteins to mediate endonucleolytic cleavage of target RNAs[17].

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 two 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 pro-differentiating *TINCR*[24]. A recent report defined very long intervening ncRNAs (up to 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], though 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 (NATs) 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 NATs 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 over 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 (SRF)[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 a number of EC-specific genes[42]. Normal differentiated properties of SMC and EC further require fine-tuning of gene expression through the action of microRNAs[43]. Since 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 HCASMC as a first step towards understanding the potential role of lncRNAs in human SMC phenotypic control. Here, we report on the identification of 31 lncRNAs, including one named *SENCR* (for Smooth muscle and Endothelial cell enriched migration/differentiation-associated long Non-Coding RNA). We have characterized the expression, splicing, and localization of *SENCR* and have identified unique gene signatures upon its knockdown in SMC. *SENCR* appears 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). 79.41% of filtered reads could be aligned to the human reference genome. 31 lncRNAs met our strict inclusion criteria (Methods) with the majority (22/31) falling into the lincRNA subclass (Table II in the online only Data supplement). Conventional RT-PCR showed detectable expression of 21/31 lncRNAs in a panel of human cell types, including HCASMC and 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 plasma (Figure 1B–1C). One of the lncRNAs (*lncRNA9*) exhibited a selective pattern of expression in cell lines (Figure 1A, 1D) and human tissues (Figure 1A, 1B). We refer to this lncRNA as *SENCR* because of its enriched expression in both smooth muscle and endothelial cells (Figure 1A, 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 three 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 endothelial cell and blood cell formation[50] (Figure 2A). There is no overlap between *SENCR* and *FLI1* exonic sequences indicating *SENCR* is not a natural anti-sense 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 this transcript has no or low protein-coding potential (not shown). Primers to exons 1 and 3 of *SENCR* showed the presence of two 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, 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, 2D). In general, there was coincident expression of *SENCR* with *FLI1* suggesting these transcripts may be under similar transcriptional control processes (Figure 2E, 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 three 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 ENCODE data on the UCSC Genome Browser (http://genome.ucsc.edu/) support 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 towards delineating *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, though 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 FISH[51] in these two 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). On the other hand, 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 though 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 and Figure III in the online only Data supplement). We next used two distinct probe sets to *SENCR* in HUVEC treated with a control dicer substrate RNA or two 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 upon *SENCR* knockdown (Figure 4D). These results establish the cytoplasmic localization of *SENCR* and indicate its relatively weaker level of expression as compared to 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 upon knocking down *SENCR* with multiple dicer substrate RNAs in HCASMC (Figure 5A–5C) or HUVEC (Figure 5D–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, 6H). Further, 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 towards understanding the function of *SENCR*, we performed RNA-seq in HCASMC following knockdown of *SENCR* to assess changes in the transcriptome. Most sequencing reads were aligned to the reference genome and scatterplots of replicates showed very similar transcript profiles (not shown). Statistical analysis of each set of replicates revealed hundreds of genes that were significantly induced or repressed upon *SENCR* knockdown (Figure 6A and Table III in the online only Data supplement). Strikingly, many SMC contractile genes showed significant reduction in mRNA expression with *SENCR* knockdown (Figure 6B and Table 3 in the online only Data supplement). Gene ontology analysis using DAVID revealed biological processes associated with this reduced contractile gene signature (Table 4 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 down-regulation in HCASMC (Figure 6D). We also confirmed reduced expression of several of the SMC contractile genes at both the mRNA level (Figure 6E) and protein level (Figure 6G). While the SMC contractile program was reduced with *SENCR* knockdown, a number of genes associated with cell migration were induced (Figure 6C and Table 3 in the online only Data supplement). DAVID analysis supported biological processes linked to cellular locomotion with *SENCR* knockdown (Table 5 in the online only Data supplement). We validated two 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 pro-migratory gene signature.

Attenuated *SENCR* expression confers a hyper-motile phenotype in HCASMC

To ascertain whether the increase in pro-migratory gene expression upon knockdown of *SENCR* translates into a functional phenotype, we performed two independent measures of cell migration. Using a scratch wound assay, we observed hyper-motile HCASMC with *SENCR* knockdown (Figure 7A, 7B). Many of these cells exhibited reorganization of the actin cytoskeleton with formation of lamellipodia, consistent with a migratory cell phenotype (arrows in Figure 7Af). Importantly, the increase in HCASMC migration could be completely rescued upon simultaneous knockdown of either of two pro-migratory genes shown to be induced upon knockdown of *SENCR* (Figure 7C and Figure V in the online only Data supplement). To further confirm this accentuated cell migration phenotype upon 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 though not quite as much as that observed with the potent migratory stimulus, PDGF-BB (Figure 7D, 7E). We also observed augmented PDGF-BB-induced cell migration upon 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 long noncoding RNA (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 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 endothelial cells. 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 appear to regulate local gene expression in *cis*. Indeed, our extensive RNA FISH and biochemical fractionation studies clearly indicate *SENCR* to be a cytoplasmic lncRNA supporting an extranuclear function. Using RNA-seq following knockdown of *SENCR*, we observed uniform decreases in expression of SMC contractile-associated genes as well as attenuated expression of the major transcriptional switch (Myocardin) for the differentiation of vascular SMC[39]. On the other hand, knockdown of *SENCR* augments a pro-migratory gene signature that facilitates heightened SMC migration. Thus, we have uncovered a new vascular cell-enriched lncRNA that appears to function in the maintenance of a normal, non-motile 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 NATs (47.0%), followed by intronic (18.8%), divergent (16.4%), completely overlapping (7.4%), 5' overlapping (7.1%) and 3' overlapping (3.4%) lncRNAs (Supplemental Table 6). Much of what is known about sense-antisense gene pairs relates to NATs 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* appear to be co-expressed 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 is 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 though an endothelial cell-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 very 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 bi-directionally transcribed protein-coding genes[61]. However, it is completely unclear how other sense-antisense pairs may be transcribed, particularly a lncRNA that is co-expressed 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, though the level of activity remained much lower when compared to an endothelial cell-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 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 owing to 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 pro-migratory genes leading to enhanced cell motility. The mechanism for such changes in cell phenotype is unknown at this time; however, since *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, de-stabilization 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 ~1,000 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. On the other hand, mouse-specific lncRNAs may have limited translational relevenace 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 in order 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 kcal/mol 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 (*e.g.*, 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 6 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 endothelial cell development as well as disease-associated processes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors 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.

Source of funding

This work was supported by grants from the Nationals Institutes of Health (HL62572 and HL091168 to JMM; MH099452 to DZ and partially by HL111770; and 5P01-CA013106 to DLS) and the American Heart Association (10SDG3670036 to XL and 12POST11950002 to RDB). JHB was funded by a DAAD postdoctoral fellowship.

Abbreviations

ENCODE	Encyclopedia of DNA elements
FLI1	Friend leukemia virus integration 1
HCASMC	human coronary artery smooth muscle cell(s)
HUVEC	human umbilical vein endothelial cell(s)
lncRNA	long noncoding RNA
MYOCD	myocardin
RNA FISH	RNA fluorescence in situ hybridization
SENCR	smooth muscle and endothelial cell enriched migration/differentiation-associated long Non-Coding RNA

References

1. Ohno S. So much "junk" DNA in our genome. Brookhaven Symp Biol. 1972; 23:366–370. [PubMed: 5065367]
2. Okazaki Y, Furuno M, Kasukawa T, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cdnas. Nature. 2002; 420:563–573. [PubMed: 12466851]
3. Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SPA, Gingeras TR. Large-scale transcriptional activity in chromosomes 21 and 22. Science. 2002; 296:916–919. [PubMed: 11988577]
4. Carninci P, Kasukawa T, Katayama S, et al. The transcriptional landscape of the mammalian genome. Science. 2005; 309:1559–1563. [PubMed: 16141072]
5. Cheng J, Kapranov P, Drenkow J, et al. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science. 2005; 308:1149–1154. [PubMed: 15790807]

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. [PubMed: 22081020]
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 research. 2013; 41:6300–6315. [PubMed: 23630323]
8. Gerstein MB, Kundaje A, Hariharan M, et al. Architecture of the human regulatory network derived from ENCODE data. Nature. 2012; 489:91–100. [PubMed: 22955619]
9. Neph S, Vierstra J, Stergachis AB, et al. An expansive human regulatory lexicon encoded in transcription factor footprints. Nature. 2012; 489:83–90. [PubMed: 22955618]
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. [PubMed: 15753290]
11. Benson CC, Zhou Q, Long X, Miano JM. Identifying functional single nucleotide polymorphisms in the human CArGome. Physiol Genomics. 2011; 43:1038–1048. [PubMed: 21771879]
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 Genetics. 2012; 8:e1002624. [PubMed: 22496669]
13. The Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]
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. [PubMed: 23172862]
15. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science. 2001; 294:853–858. [PubMed: 11679670]
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. [PubMed: 23180440]
17. Ghildiyal M, Zamore PD. Small silencing RNAs: An expanding universe. Nature Reviews Genetics. 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. [PubMed: 19182780]
19. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea MD, 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. [PubMed: 19571010]
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. [PubMed: 21890647]
21. Lee JT. Epigenetic regulation by long noncoding RNAs. Science. 2012; 338:1435–1439. [PubMed: 23239728]
22. Ulitsky I, Bartel DP. lincRNAs: Genomics, evolution, and mechanisms. Cell. 2013; 154:26–46. [PubMed: 23827673]
23. Gutschner T, Hammerle M, Diederichs S. MALAT1 -- a paradigm for long noncoding RNA function in cancer. J Mol Med (Berl). 2013; 91:791–801. [PubMed: 23529762]
24. Kretz M, Siprashvili Z, Chu C, et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature. 2012; 493:231–235. [PubMed: 23201690]
25. St Laurent G 3rd, 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. [PubMed: 23876380]
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. [PubMed: 20485488]

27. Orom 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. [PubMed: 20887892]
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. [PubMed: 22078386]
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. [PubMed: 23006825]
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 Research. 2004; 32:4812–4820. [PubMed: 15356298]
31. Katayama S, Tomaru Y, Kasukawa T, et al. Antisense transcription in the mammalian transcriptome. Science. 2005; 309:1564–1566. [PubMed: 16141073]
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 Research. 2012; 22:1775–1789. [PubMed: 22955988]
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. [PubMed: 22541732]
34. Lee JT, Bartolomei MS. X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell. 2013; 152:1308–1132. [PubMed: 23498939]
35. Pasmant E, Sabbagh A, Vidaud M, Bieche I. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. FASEB J. 2011; 25:444–448. [PubMed: 20956613]
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 Research. 2013; 41:D246–D251. [PubMed: 23042674]
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. [PubMed: 3203386]
38. Wang DZ, 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. [PubMed: 11439182]
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. [PubMed: 12392995]
40. Kume T. Foxc2 transcription factor: A newly described regulator of angiogenesis. Trends Cardiovasc Med. 2008; 18:224–228. [PubMed: 19185813]
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. [PubMed: 22406820]
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. [PubMed: 19070576]
43. Small EM, Olson EN. Pervasive roles of microRNAs in cardiovascular biology. Nature. 2011; 469:336–342. [PubMed: 21248840]
44. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012; 81:145–166. [PubMed: 22663078]
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. [PubMed: 8636440]
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. [PubMed: 19880500]

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 Comm. 2012; 419:612–616. [PubMed: 22382030]
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. [PubMed: 22706276]
49. Leung A, Trac C, Jin W, Lanting L, Akbany A, Saetrom 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. [PubMed: 23697773]
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. [PubMed: 18718762]
51. 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. [PubMed: 24097269]
52. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009; 136:629–641. [PubMed: 19239885]
53. Fatica A, Bozzoni I. Long non-coding RNAs: New players in cell differentiation and development. Nature Reviews Genetics. 2013; 15:7–21.
54. Batista PJ, Chang HY. Long noncoding RNAs: Cellular address codes in development and disease. Cell. 2013; 152:1298–1307. [PubMed: 23498938]
55. Kung JT, Colognori D, Lee JT. Long noncoding RNAs: Past, present, and future. Genetics. 2013; 193:651–669. [PubMed: 23463798]
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. [PubMed: 23898077]
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. [PubMed: 18978789]
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. [PubMed: 12045153]
59. Lapidot M, Pilpel Y. Genome-wide natural antisense transcription: Coupling its regulation to its different regulatory mechanisms. EMBO reports. 2006; 7:1216–1222. [PubMed: 17139297]
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. [PubMed: 20228226]
61. Wakano C, Byun JS, Di LJ, Gardner K. The dual lives of bidirectional promoters. Biochimica Biophysica 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. [PubMed: 23041286]
63. Barbeau B, Bergeron D, Beaulieu M, Nadjem Z, Rassart E. Characterization of the human and mouse Fli-1 promoter regions. Biochimica Biophysica Acta. 1996; 1307:220–232.
64. Raslova H, Komura E, Le Couedic 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. [PubMed: 15232614]
65. Deng Q, Ramskold D, Reinius B, Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science. 2014; 343:193–196. [PubMed: 24408435]
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. [PubMed: 17241883]
67. Yang L, Froberg JE, Lee JT. Long noncoding RNAs: Fresh perspectives into the RNA world. Trends Biochem Sci. 2013; 39:35–43. [PubMed: 24290031]

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. [PubMed: 22000014]
69. Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, Finch CE, St LG III, 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. [PubMed: 18587408]
70. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. Nature. 2011; 470:284–288. [PubMed: 21307942]
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. [PubMed: 23064229]
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. [PubMed: 24091684]
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. [PubMed: 22196729]
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. [PubMed: 16751343]
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. [PubMed: 23643243]
76. Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012; 337:1190–1195. [PubMed: 22955828]

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

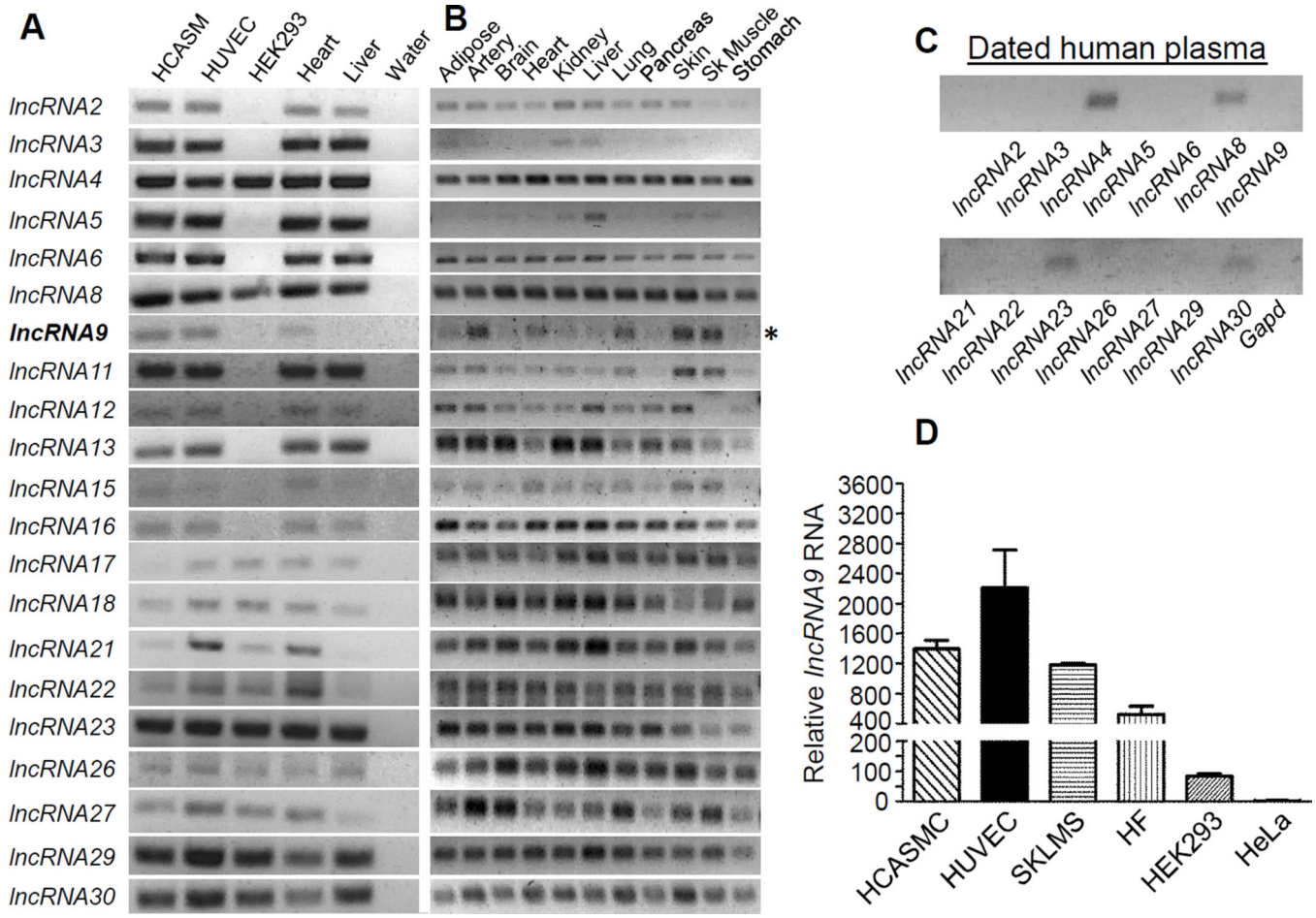


Figure 1. Validation of lncRNA expression in human cells and tissues

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.

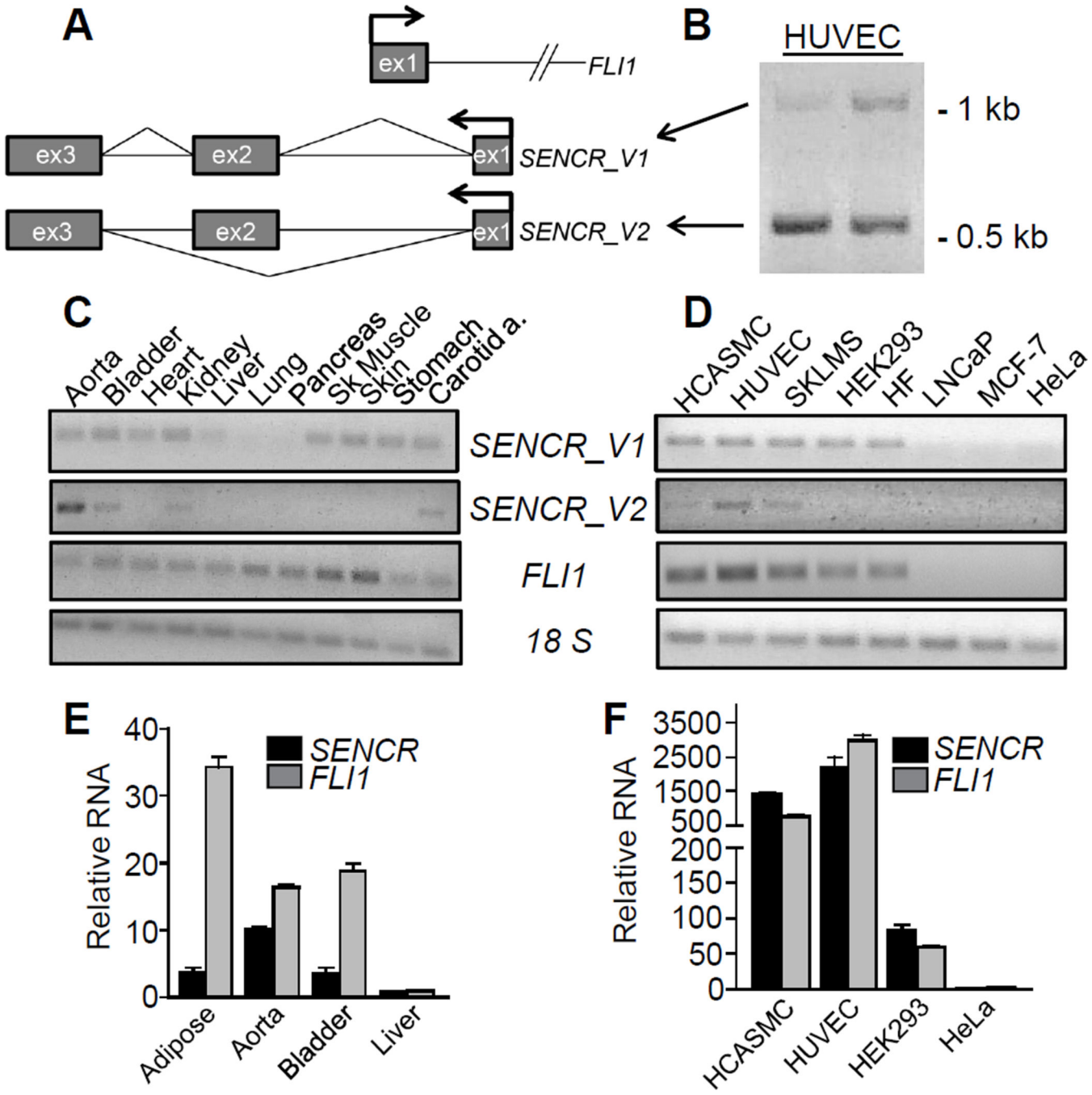


Figure 2. *SENCR* gene structure and isoform expression

(**A**) Schematic of *SENCR* and *FLI1* (partial) gene loci. Arrows denote the transcription start sites and bent lines in *SENCR* indicate splicing patterns. (**B**) RT-PCR of *SENCR* with primers to exon 1 and 3 showing the presence of two transcripts reflecting full length (V1) and alternately spliced (V2) *SENCR*. RT-PCR of two *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 standard deviation of one experiment with three biological replicates. All expression data here and below

represent at least two (more typically multiple) independent studies performed by more than one author. Unless indicated otherwise, *SENCR* expression here and below reflects both isoforms using primers to a common exon. SKLMS, uterine leiomyosarcoma cell line; HF, human fibroblasts.

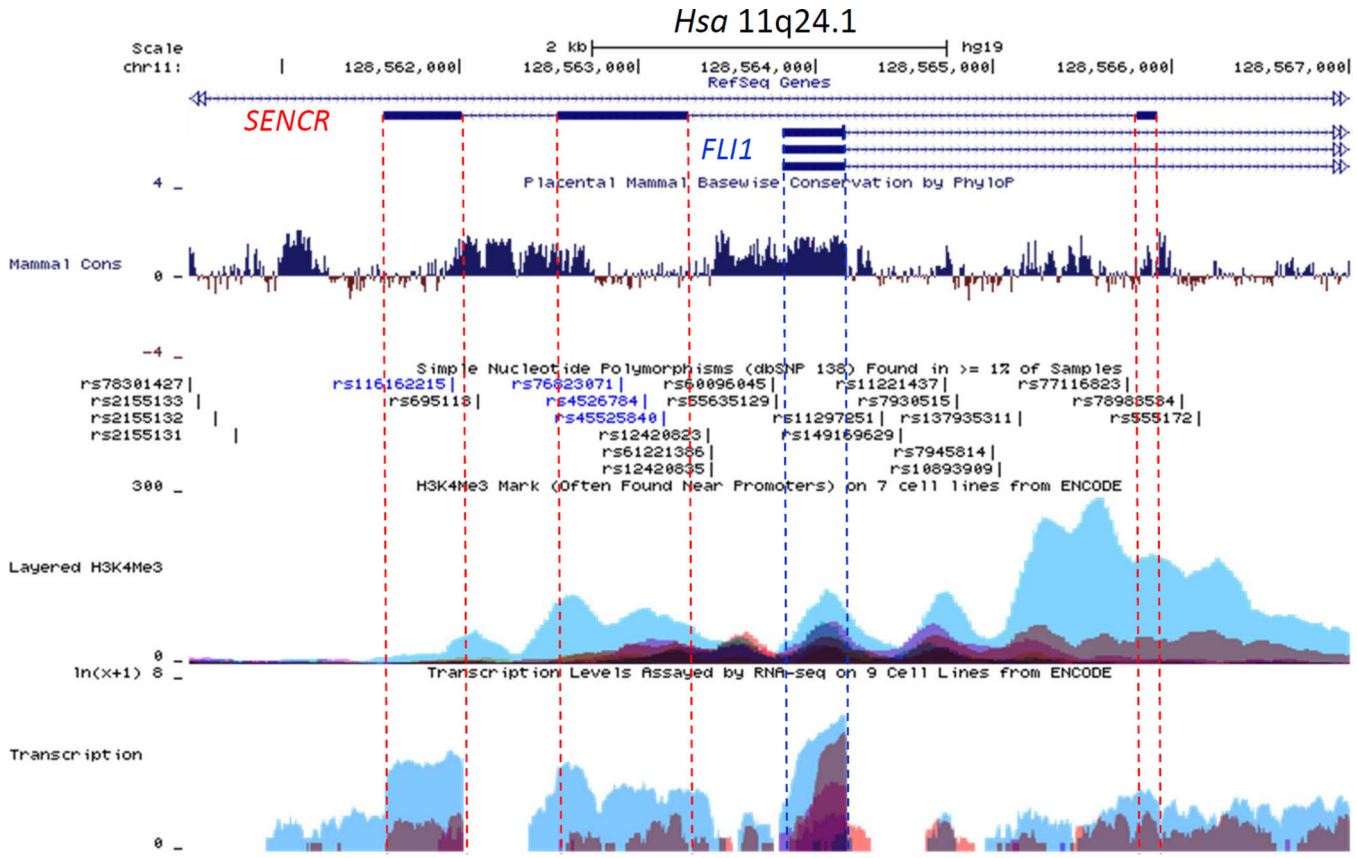


Figure 3. UCSC genome browser track of the human *FLI1-SENCR* sense-antisense gene pair

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 ENCODE. Note the lower conservation and selective transcription in HUVEC (blue peaks at bottom) for *SENCR* as compared to *FLI1*. See Results for more details.

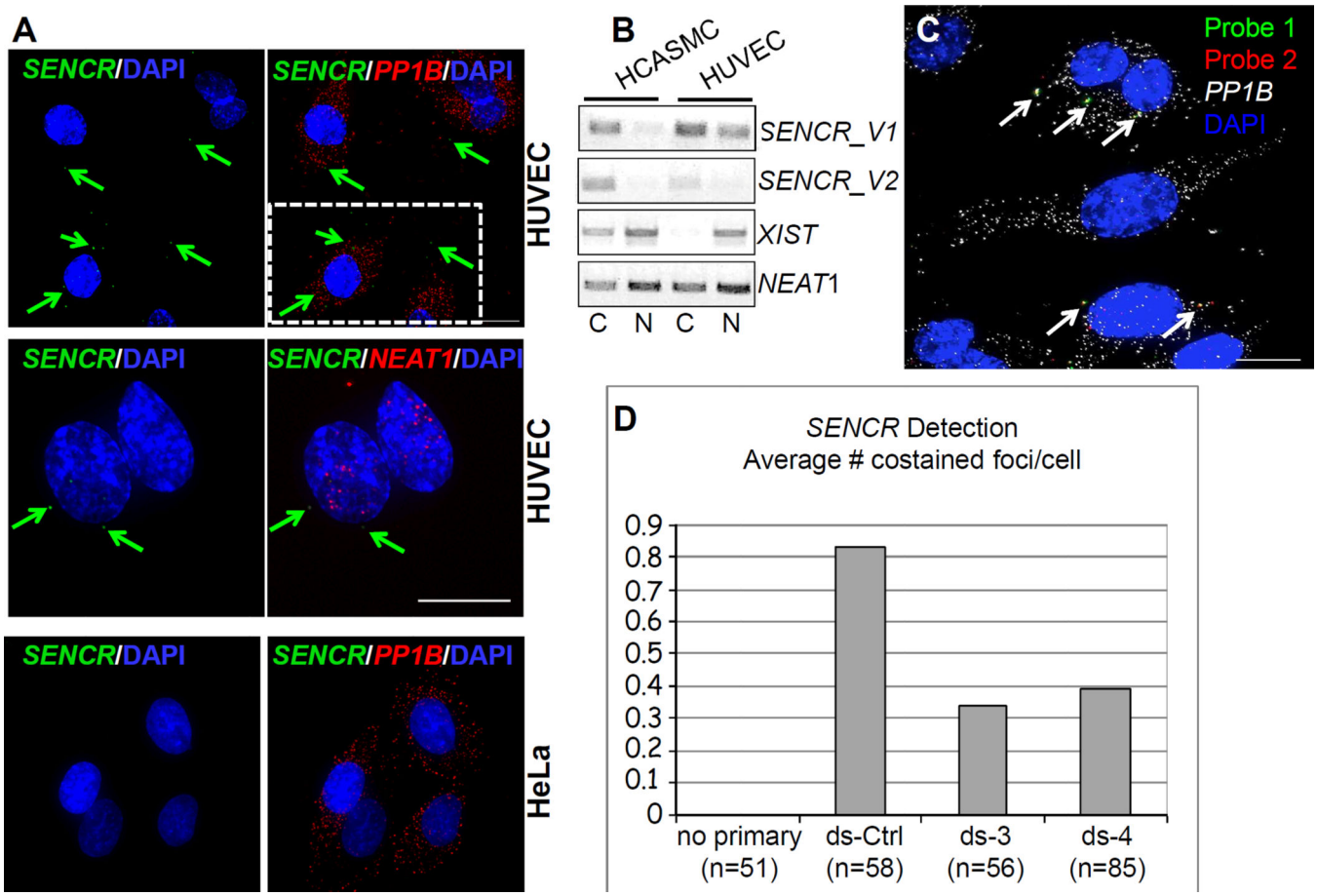


Figure 4. Localization of *SENCR* RNA

(**A**) RNA FISH analysis of *SENCR* versus cytoplasmic *PP1B* mRNA and nuclear *NEAT1* RNA in indicated cell types. Arrows point to single molecules of *SENCR* RNA in the cytoplasm of 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**) RT-PCR analysis of two *SENCR* isoforms versus other lncRNA genes from polyA+ RNA isolated from the cytoplasmic (C) or nuclear (N) fractions of indicated cell types. (**C**) Application of two probe sets to *SENCR* RNA (see Methods). Arrows point to coincident localization of two 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 two dicer substrate RNAs targeting different regions of *SENCR* (ds-3 and ds-4). The Y-axis indicates the average number of co-stained foci/cell.

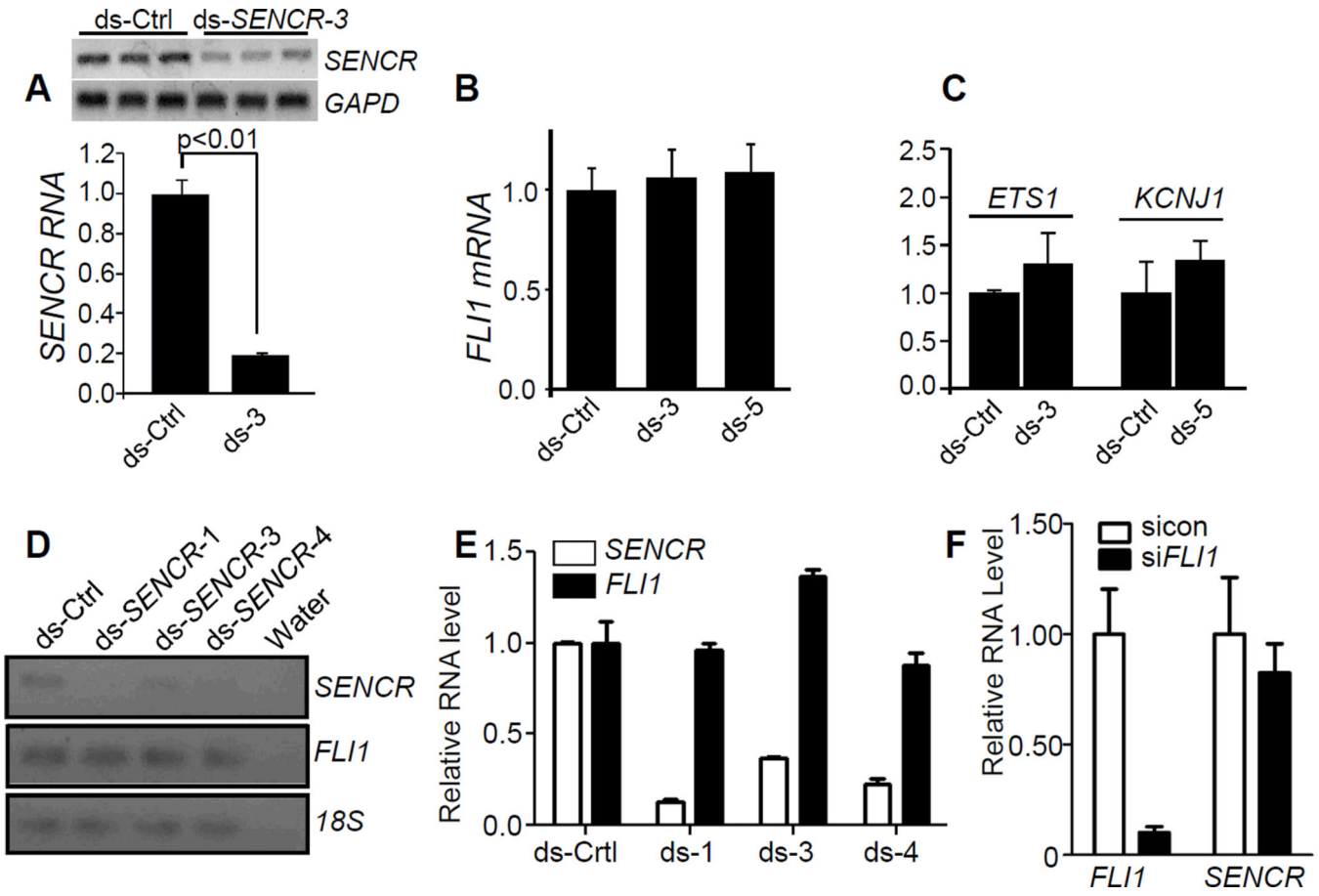
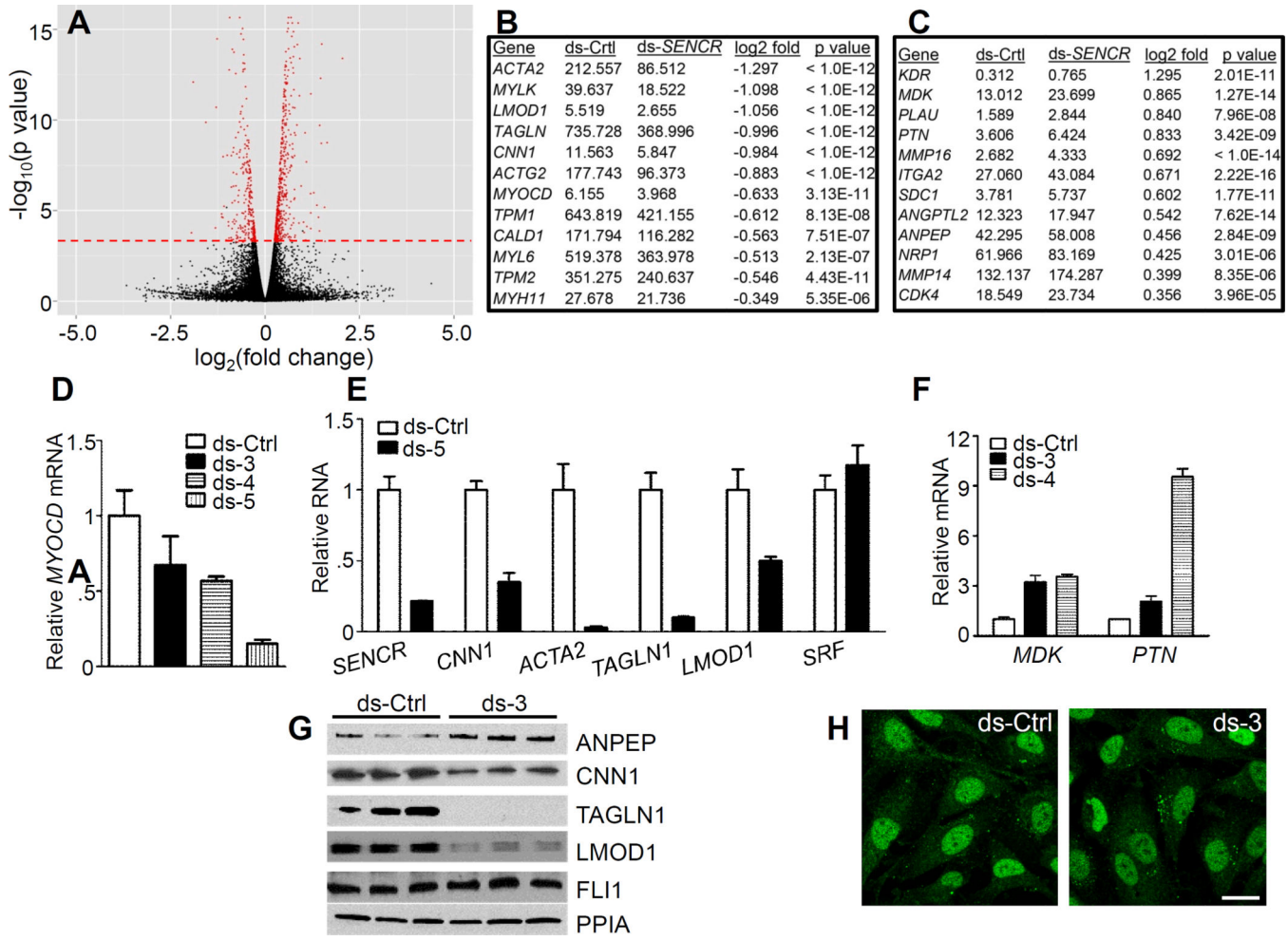


Figure 5. Effect of knocking down *SENCR* on local gene expression

(**A**) Dicer substrate control (ds-Ctrl) or dicer substrate *SENCR* RNA was transfected into HCASMC for 72 hr and then total RNA isolated for conventional (top) or quantitative (bottom) 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**) following 3 d transfection with indicated dicer substrate RNAs. (**D**) Conventional RT-PCR of *SENCR* and *FLI1* in HUVEC following transfection with indicated dsRNAs. (**E**) Quantitative RT-PCR of *SENCR* and *FLI1* following transfection with indicated dsRNAs in HUVEC. (**F**) Quantitative RT-PCR of *FLI1* and *SENCR* following knockdown of *FLI1* mRNA in HCASMC. Data are representative of multiple independent experiments carried out 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 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 SMC contractile genes (**B**) and pro-migratory genes (**C**) exhibiting either reduced (**B**) or increased (**C**) expression with ds-*SENCR* knockdown. See Table 3 in the online only Data supplement for a complete listing of all genes showing significant up- or down-regulation with *SENCR* knockdown. Quantitative RT-PCR validation of reduced *MYOCD* mRNA (**D**) and SMC contractile genes (**E**) in HCASMC following knockdown of *SENCR* with various dsRNAs. (**F**) Quantitative RT-PCR validation of up-regulation of two pro-migratory genes upon knockdown of *SENCR* in HCASMC. (**G**) Western blot validation of up-regulated (ANPEP) and down-regulated (SMC contractile) proteins in HCASMC 72 hr after indicated transfection with dsRNA. Similar findings were observed in an independent experiment. (**H**) Immunofluorescence microscopy of FLI1 protein in the nucleus of HCASMC following 3 d transfection with indicated dsRNAs. Results are representative of multiple experiments performed by independent investigators. Scale bar = 10 μm for both images.

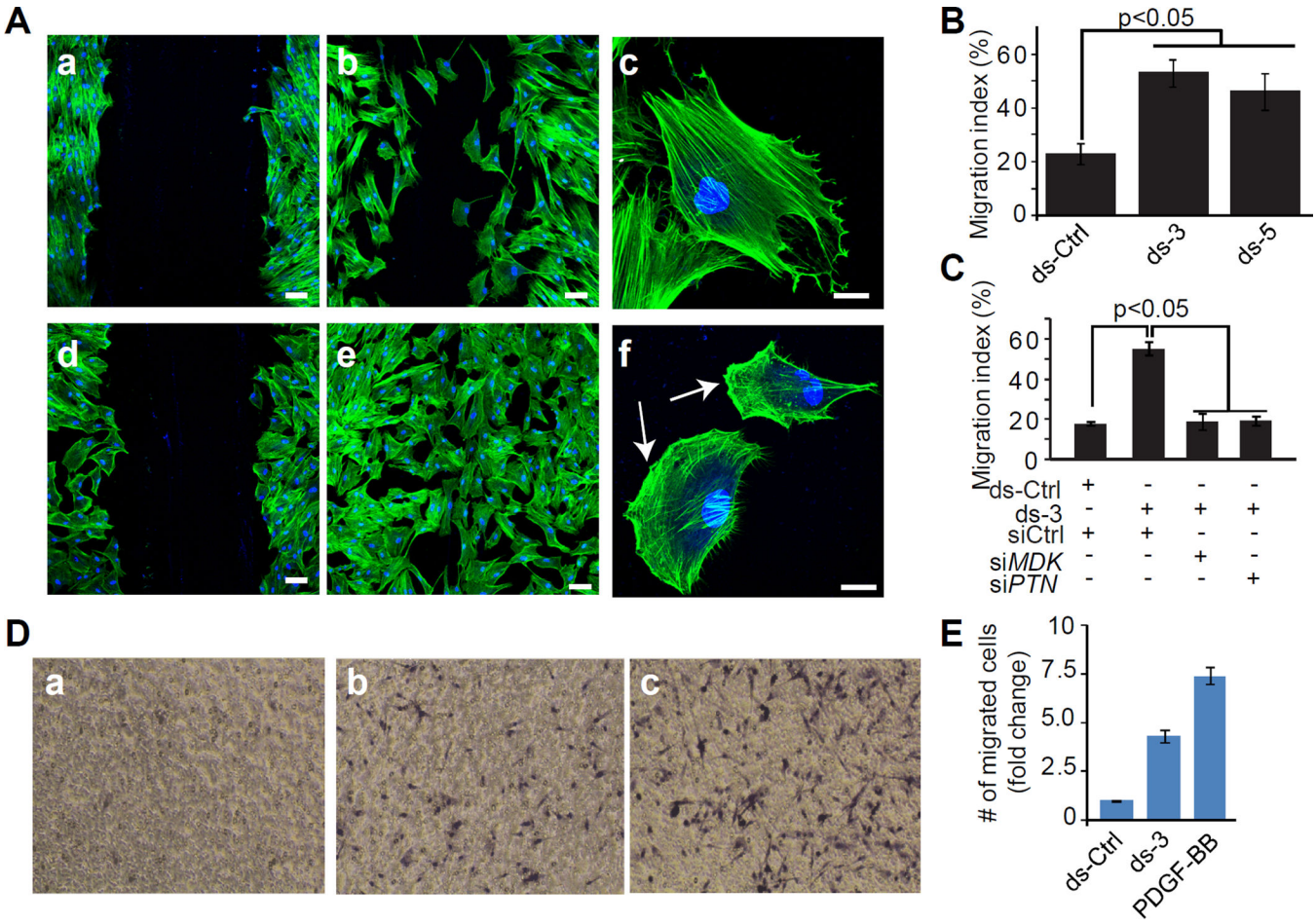


Figure 7. Effect of *SENCR* knockdown in a scratch wound assay of cell migration

(**A**) HCASMC were transfected with ds-Ctrl (panels a–c) or ds-*SENCR*-3 (panels d-f) for 72 hr after which a scratch wound was created and cell migration assessed in quiescent (panels a, d) HCASMC or in similar cells stimulated for 12 hr with 10% FBS (panels b, c, e, f). Cells were stained with phalloidin (red) and DAPI (blue). Bars are 25 μm in panels a, b, d, and e and 10 μm in panels c and f. Arrows in panel f denote lamellipodia. (**B**) Quantitative measure of the area of the wound occupied by dsRNA-transfected HCASMC 12 hrs following serum stimulation. (**C**) Same experiment as in B only HCASMC were transfected simultaneously with a control siRNA (siCtrl) or an siRNA to one of two pro-migratory genes. Each siRNA reduced level of *MDK* or *PTN* mRNA by more than 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 hr and the fold change in number of cells migrating through the porous membrane quantitated (**E**). The data reflect cell counts from 5 independent fields.