

RESEARCH ARTICLE

Identification of long noncoding RNAs involved in muscle differentiation

Yeong-Hwan Lim^{1,2,3}, Duk-Hwa Kwon^{1,4}, Jaetaek Kim^{1,5}, Woo Jin Park^{1,6}, Hyun Kook^{1,3,4}, Young-Kook Kim^{1,2,3*}

1 Basic Research Laboratory for Cardiac Remodeling Research Laboratory, Chonnam National University Medical School, Jeollanam-do, Republic of Korea, **2** Department of Biochemistry, Chonnam National University Medical School, Jeollanam-do, Republic of Korea, **3** Department of Biomedical Sciences, Center for Creative Biomedical Scientists at Chonnam National University, Jeollanam-do, Republic of Korea, **4** Department of Pharmacology, Chonnam National University Medical School, Jeollanam-do, Republic of Korea, **5** Division of Endocrinology and Metabolism, Department of Internal Medicine, College of Medicine, Chung-Ang University, Seoul, Republic of Korea, **6** College of Life Sciences, Gwangju Institute of Science and Technology (GIST), Gwangju, Republic of Korea

* ykk@jnu.ac.kr



OPEN ACCESS

Citation: Lim Y-H, Kwon D-H, Kim J, Park WJ, Kook H, Kim Y-K (2018) Identification of long noncoding RNAs involved in muscle differentiation. *PLoS ONE* 13(3): e0193898. <https://doi.org/10.1371/journal.pone.0193898>

Editor: Seungil Ro, University of Nevada School of Medicine, UNITED STATES

Received: September 28, 2017

Accepted: February 19, 2018

Published: March 2, 2018

Copyright: © 2018 Lim et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This research was financially supported by the grants from the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A02036313, NRF-2016R1A4A1009895 to YKK). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Long noncoding RNAs (lncRNAs) are a large class of regulatory RNAs with diverse roles in cellular processes. Thousands of lncRNAs have been discovered; however, their roles in the regulation of muscle differentiation are unclear because no comprehensive analysis of lncRNAs during this process has been performed. In the present study, by combining diverse RNA sequencing datasets obtained from public database, we discovered lncRNAs that could behave as regulators in the differentiation of smooth or skeletal muscle cells. These analyses confirmed the roles of previously reported lncRNAs in this process. Moreover, we discovered dozens of novel lncRNAs whose expression patterns suggested their possible involvement in the phenotypic switch of vascular smooth muscle cells. The comparison of lncRNA expression change suggested that many lncRNAs have common roles during the differentiation of smooth and skeletal muscles, while some lncRNAs may have opposite roles in this process. The expression change of lncRNAs was highly correlated with that of their neighboring genes, suggesting that they may function as cis-acting lncRNAs. Furthermore, within the lncRNA sequences, there were binding sites for miRNAs with expression levels inversely correlated with the expression of corresponding lncRNAs during differentiation, suggesting a possible role of these lncRNAs as competing endogenous RNAs. The lncRNAs identified in this study will be a useful resource for future studies of gene regulation during muscle differentiation.

Introduction

The major type of smooth muscle cells, vascular smooth muscle cells (VSMCs), are capable of converting between synthetic and contractile phenotypes [1, 2]. In the normal state, VSMCs exist as the differentiated and contractile type. However, in response to injury, VSMCs become de-differentiated into a more proliferative and synthetic phenotype. Diverse factors and complex

Competing interests: The authors have declared that no competing interests exist.

regulatory networks govern the phenotypic change of VSMCs [3]. Disturbances in the proliferation or differentiation of VSMCs can lead to various vascular diseases, such as atherosclerosis, restenosis, and hypertension, and intensive research is underway to identify the diverse factors involved in this process [4].

In contrast to the reversible change in phenotype of smooth muscle cells, terminally differentiated skeletal muscle, like many other cell types, cannot be dedifferentiated. Upon external signals such as muscle injury, skeletal muscle is regenerated from satellite cells, which are main myogenic progenitor cells. The satellite cells are divided asymmetrically, and expanded into myoblasts, which are further differentiated into myocytes. This differentiation process is influenced by diverse factors, and for the treatment of diseases such as skeletal muscle degeneration, the elucidation of signaling networks involved in this process is critical [5, 6].

Many studies have shown that various types of noncoding RNAs have regulatory roles in most cellular processes. MicroRNAs (miRNAs), a class of small noncoding RNAs, have been studied extensively for a long time [7]. The regulatory roles of dozens of miRNAs have been reported for diverse aspects of smooth muscle cells [8]. Long noncoding RNAs (lncRNAs) are another group of regulatory RNAs that are longer than 200 nucleotides but lack protein-coding potential [9]. Compared with miRNAs, only a few studies have identified the roles of lncRNAs for smooth muscle, including the regulation of proliferation or apoptosis. One of the first identified lncRNAs was the smooth muscle and endothelial cell-enriched migration/differentiation-associated lncRNA (*SENCR*) [10]. Inhibition of *SENCR* increased the migration of smooth muscle cells. Another lncRNA essential for the proliferation of VSMCs is smooth muscle-induced lncRNA enhances replication (*SMILR*) [11]. *SMILR* regulates its neighboring protein-coding gene, hyaluronan synthase 2 (*HAS2*), an enzyme that synthesizes hyaluronic acid. The function of only a small number of lncRNAs has been elucidated in the phenotypic change of VSMCs; therefore, it is hypothesized that additional lncRNAs involved in this process await discovery. To date, various reagents have been used to model the differentiation process of VSMCs, such as platelet-derived growth factor (PDGF), myocardin (MYOCD), transforming growth factor beta (TGF β), activin A, retinoids, and angiotensin II. The transcriptome responds quite differentially to each of these factors; therefore, to identify common lncRNAs, diverse experimental models should be combined.

In the present study, we aimed to comprehensively discover lncRNAs involved in muscle differentiation by combining diverse RNA sequencing (RNA-seq) datasets. By integrating the data from VSMCs where different extracellular cues were used for differentiation, we identified several lncRNAs with potential as regulators in VSMCs differentiation. We analyzed the correlation between the expression of lncRNAs and that of their neighboring protein-coding genes, and compared the expression change of lncRNAs between smooth and skeletal muscle cells.

Materials and methods

Analysis of RNA-seq data

We obtained RNA-seq data from the Gene Expression Omnibus (GEO) database [12]. These data included RNA-seq data from platelet-derived growth factor (PDGF)-treated venous smooth muscle cells (GSE69637), myocardin (MYOCD)-overexpressed human coronary artery smooth muscle (HCASM) cells (GSE77120), transforming growth factor beta (TGF β)-treated HCASM cells (GSE85910), and human myoblasts differentiated with low serum (GSE70389). The Trimmomatic algorithm was used to remove adaptor sequences and filter the reads with low quality (option: ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15) [13]. Reads were aligned to the human genome

(hg19) using STAR with default options [14]. SAMtools was used to remove reads with low mapping quality (option: -q 10) [15]. To calculate the FPKM (fragments per kilobase of exon model per million mapped fragments), Cuffnorm was used together with the GENCODE annotation for lncRNAs (GRCh37, release 26; <http://www.genencodegenes.org/>) [16], and the annotation of UCSC genes for protein-coding genes (hg19; <http://genome.ucsc.edu/>) [17]. The GENCODE annotation contains information for 15,901 lncRNAs. We removed the lncRNAs with an average FPKM value in each dataset lower than 5. This filtering resulted in 2,361 lncRNAs remaining for further analysis.

Selection of candidate lncRNAs involved in muscle differentiation

To select the lncRNAs with differential expression during differentiation of VSMCs, we selected only those lncRNAs with a log₂ value of expression changes greater than 0.5 or lower than -0.5 between before and after treatment with external factors. To select the lncRNAs that increased in the synthetic phenotype, those lncRNAs with increased expression in the PDGF-treated sample, but with decreased expression in the MYOCD- and TGFβ-treated samples, were chosen. In the case of the lncRNAs that increased in the contractile phenotype, we chose those lncRNAs with decreased expression in the PDGF-treated sample, but with increased expression in the MYOCD- and TGFβ-treated samples. We removed the lncRNAs from the list if the FPKM values of $\geq 50\%$ of the samples belonging to the denominator were zero. To select statistically significant lncRNAs, a two-tailed T-test was applied to each dataset, and labeled as 'significant' if the *p* values were less than 0.1 in more than two datasets and less than 0.3 in all datasets.

To select lncRNAs that are differentially expressed during skeletal muscle differentiation (GSE70389), we used the same criteria that were used for lncRNA selection in VSMCs, except that the *p* value criteria in the T-test as 0.2.

In silico promoter analysis

To predict transcription factors that might regulate the expression of lncRNAs, we used the same approach as that detailed in our previous report for the analysis of miRNA promoters [18]. We used the data released by the ENCODE project, where chromatin immunoprecipitation followed by sequencing (ChIP-seq) is performed for 161 transcription factors [19]. We selected the genomic region spanning from -2,000 to +500 nucleotides against the 5' end of the lncRNA as the promoter of the lncRNA. For the lncRNAs with incomplete annotation in GENCODE, we modified the annotation by comparison with the information from RefSeq. After collecting the genomic coordinates of the lncRNA promoters, we searched for ChIP-seq signals of transcription factors within this region.

Measurement of RNAs during the differentiation of HCASM

To measure the expression level of lncRNAs and their neighboring protein-coding genes during differentiation of smooth muscle cells, we cultured HCASM cells (Gibco) according to the manufacturer's protocol. These cells were maintained in Medium 231 (Gibco) supplemented with Smooth Muscle Growth Supplement (Gibco), and harvested for the sample of synthetic phenotype. To obtain the sample of contractile or differentiated phenotype, HCASM cells stabilized in the growth media were cultured in Medium 231 supplemented with Smooth Muscle Differentiation Supplement (Gibco) for three days, and then collected. We used TRIzol Reagent (Invitrogen) to extract total RNA, performed reverse transcription using RevertAid Reverse Transcriptase (Thermo Scientific) with Random Hexamers (Invitrogen). Quantitative real time polymerase chain reaction (PCR) was performed using SYBR Green PCR Master

Mix (Applied biosystems) in Rotor-Gene Q (Qiagen). The primer sequences for PCR were included in the supplementary table. The differentiation of these cells was confirmed by measuring the level of Calponin 1 (CNN1) and Smooth muscle protein 22-alpha (SM22 α). For normalization, the expression level of Glyceraldehyde 3-phosphate dehydrogenase (GADPH) was used.

Analysis of lncRNAs as miRNA regulators

In the track data hubs of UCSC Genome Browser [17], we utilized miRcode, which is a prediction algorithm of human miRNA targets based on GENCODE annotation including lncRNAs [20]. We intersected the miRcode track with lncRNA track using Table browser, to collect the list of lncRNA-targeting miRNAs. To select more reliable miRNAs, we obtained miRNA microarray data for human VSMCs undergoing phenotypic switching in response to serum withdrawal (GSE19544) [21]. Among the miRNAs predicted to target lncRNAs, we only selected the miRNAs with expression inversely correlated with lncRNAs during VSMCs differentiation.

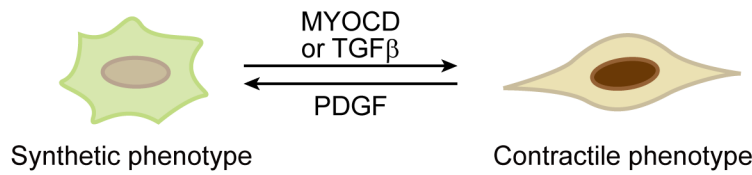
Results and Discussion

Identification of lncRNAs involved in phenotypic change of VSMCs

VSMCs can be switched between synthetic and contractile phenotypes in response to diverse extracellular cues [1, 2]. To discover lncRNAs involved in the regulation of this process, we obtained RNA-seq data from the GEO database [12]. First, RNA-seq data was obtained from VSMCs treated with PDGF and their untreated controls (GSE69637) [11]. PDGF induces the synthetic phenotype of VSMCs (Fig 1A) [1]. Data from HCASMs overexpressing myocardin (MYOCD) (GSE77120) and those cells treated with TGF β (GSE85910) were obtained [22]. Both MYOCD and TGF β were reported to induce a contractile phenotype (Fig 1A) [1, 2]. We integrated the data from artery and venous smooth muscles; therefore, it was possible to identify common lncRNAs involved in the differentiation of both types of smooth muscles. We collected these datasets and analyzed the expression levels of lncRNAs and protein-coding genes. We confirmed the phenotypic change of VSMCs from the treatment of each factor by calculating the expression level of CNN1 and smooth muscle-induced lncRNA enhances replication (SMILR) (Fig 1B). CNN1 was previously reported to increase, while SMILR is known to decrease, in the contractile phenotype of VSMCs [1, 11]. Next, we measured the expression of whole lncRNAs annotated in the recent release of the GENCODE database (release 26), and compared their expression changes during the phenotypic change of VSMCs (S1 Table). We found that several lncRNAs showed consistent differential expression between the synthetic and contractile phenotype of VSMCs in the three datasets (Fig 1C and S1 and S2 Figs). However, when we calculated the correlation for whole lncRNAs, there was almost no correlation among the three datasets (the correlation coefficients ranged from 0.02 to 0.04). Thus, most lncRNAs are not directly related with the phenotypic change of VSMCs. Otherwise, they might have a specific role during phenotypic change induced by different contexts. This result suggests that it is crucial to integrate the data from diverse datasets to identify common factors involved in a physiological process such as VSMCs differentiation.

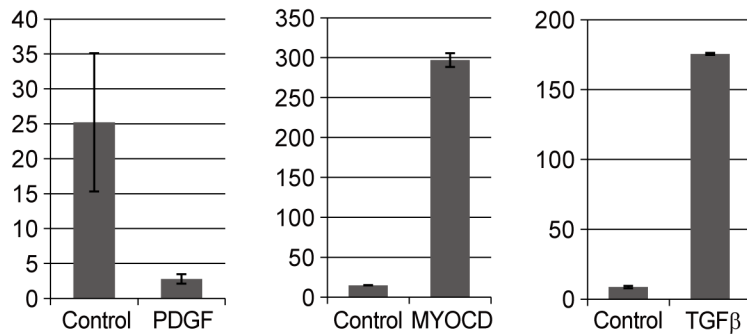
To identify those lncRNAs that are likely to play a role in phenotypic change of VSMCs, we statistically tested for the selected lncRNAs shown in S1 Table, and listed only those lncRNAs with significantly changed expression among the three datasets (Fig 2). This list included 15 lncRNAs increased in synthetic phenotype and 10 lncRNAs increased in contractile phenotype. The most highly changed lncRNAs in each direction of phenotypic conversion were SMILR and cardiac mesoderm enhancer-associated non-coding RNA (CARMN), respectively.

A Phenotypic change of VSMCs

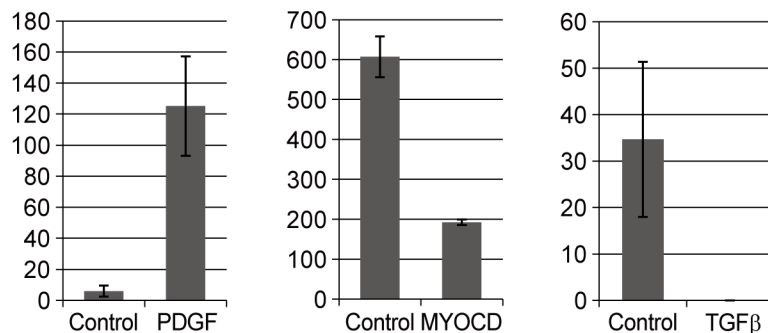


B Expression of known markers

CNN1 (Calponin 1)



SMILR (Smooth Muscle Enriched Long Noncoding RNA)



C Expression profiles of lncRNAs

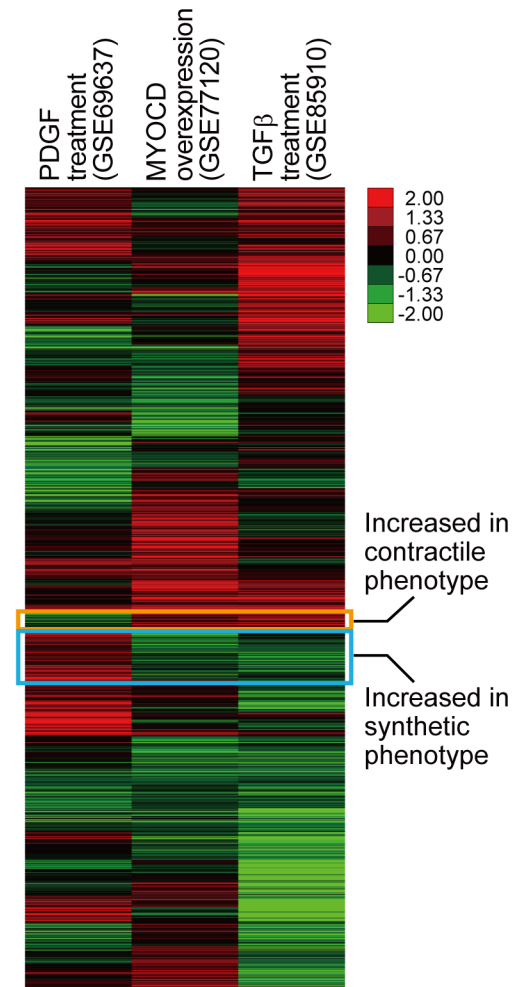


Fig 1. Identification of long noncoding RNAs (lncRNAs) involved in phenotypic change of vascular smooth muscle cells (VSMCs). (A) The VSMCs with the synthetic phenotype can be differentiated into less the proliferative and contractile phenotypic by the overexpression of myocardin (MYOCD) or by treatment with transforming growth factor beta (TGFβ). The cells with the contractile phenotype can be converted into the more proliferative and synthetic phenotype by treatment with platelet-derived growth factor (PDGF). (B) Expression level of representative genes previously known to be involved in phenotypic change of VSMCs. The RNA-seq data of PDGF-treated venous smooth muscle cells, MYOCD-overexpressing human coronary artery smooth muscle (HCASM) cells, and TGFβ-treated HCASM cells were obtained from the Gene Expression Omnibus (GEO) (see [Methods](#)). The fragments per kilobase of exon model per million mapped fragments (FPKM) values of a protein-coding gene (*CNN1*) and a long noncoding RNA (*SMILR*) are depicted in the y-axis. Error bars indicate standard errors from four samples for the PDGF set, or deviation from two samples for the MYOCD and TGFβ sets. (C) Expression profiles of lncRNAs during phenotypic change of VSMCs were analyzed. The calculation of lncRNAs' expression changes is described in Method section. The gene clusters with prominent changes between the synthetic and contractile phenotypes are indicated with boxes. Detailed expression changes of genes from those clusters are depicted in [S1](#) and [S2](#) Figs.

<https://doi.org/10.1371/journal.pone.0193898.g001>

A Selected lncRNAs with increased expression in synthetic phenotype of VSMCs

Gene ID	Gene Name	Fold(PDGF)	Fold(MYOCD)	Fold(TGFβ)
ENSG00000255364.1	SMILR	4.38	-1.66	-8.44
ENSG00000248323.5	LUCAT1	2.49	-2.89	-6.95
ENSG00000254639.1	CTD-2589M5.5	1.50	-1.49	-6.92
ENSG00000257453.1	RP11-290L1.3	2.72	-0.90	-2.50
ENSG00000260977.1	RP11-333I13.1	1.39	-1.43	-2.89
ENSG00000262211.1	CTD-2031P19.5	1.47	-1.53	-2.44
ENSG00000237499.6	RP11-356I2.4	2.07	-1.27	-1.95
ENSG00000253522.5	MIR3142HG	2.14	-1.82	-0.53
ENSG00000261468.1	RP11-1024P17.1	1.68	-0.95	-1.58
ENSG00000274021.1	RP11-823E8.3	1.75	-1.39	-0.97
ENSG00000276116.2	FUT8-AS1	0.94	-1.61	-1.30
ENSG00000245532.7	NEAT1	0.54	-0.87	-1.61
ENSG00000272711.1	RP11-259N19.1	1.15	-0.87	-0.98
ENSG00000272734.1	ADIRF-AS1	0.85	-0.57	-1.29
ENSG00000237187.8	NR2F1-AS1	0.64	-0.63	-0.56

B Selected lncRNAs with increased expression in contractile phenotype of VSMCs

Gene ID	Gene Name	Fold(PDGF)	Fold(MYOCD)	Fold(TGFβ)
ENSG00000249669.9	CARMN	-0.55	4.50	2.49
ENSG00000260604.2	RP1-140K8.5	-2.80	1.42	3.20
ENSG00000257894.2	RP1-78O14.1	-2.25	1.75	2.73
ENSG00000272970.1	RP11-329B9.4	-1.73	2.23	2.25
ENSG00000224116.6	INHBA-AS1	-2.16	2.68	1.12
ENSG00000280143.1	AP000892.6	-0.84	3.63	0.92
ENSG00000282057.1	RP4-621F18.2	-0.73	2.86	1.42
ENSG00000268129.1	RP11-91G21.1	-1.29	1.73	1.62
ENSG00000232160.6	RAP2C-AS1	-0.75	0.53	1.23
ENSG00000270005.1	RP11-380L11.3	-1.04	0.55	0.51

Fig 2. Selected long noncoding RNAs (lncRNAs) with differential expression between synthetic and contractile phenotypes of vascular smooth muscle cells (VSMCs). Based on the expression change of lncRNAs between the synthetic and contractile phenotypes from three different datasets, we selected highly reliable lncRNAs that could be involved in the regulation of phenotypic switch of VSMCs. The lncRNAs with increased expression in synthetic (A) and contractile (B) phenotypes are listed in descending order of their changes in expression, respectively. For each lncRNA, their annotated gene name and log₂(fold change) in each dataset are shown. The fragments per kilobase of exon model per million mapped fragments (FPKM) values of lncRNAs in each sample are summarized in S1 Table.

<https://doi.org/10.1371/journal.pone.0193898.g002>

SMILR was reported to be a regulator of smooth muscle cells differentiation, while the role of CARMN was identified from cardiac cells [11, 23]. As a host gene of miR-143 and miR-145, CARMN is required to maintain the differentiated state of cardiomyocytes. Thus, it may be expected that CARMN is involved in the differentiation of both smooth and cardiac muscles. The second most increased lncRNA in the synthetic phenotype of VSMCs, lung cancer

associated transcript 1 (LUCAT1), was identified as an lncRNA that increased in human pulmonary artery smooth muscle cells exposed to hypoxia, although its role in this process was not studied [24]. In addition, nuclear enriched abundant transcript 1 (NEAT1), one of the increased lncRNAs in the synthetic phenotype, has been implicated in the differentiation of skeletal muscle, as knockdown of NEAT1 increased the expression of marker genes of muscle differentiation in C2C12 cells [25]. Among the other lncRNAs in the list, *MIR3142HG* is the host gene of miR-3142 and miR-146a. miR-146a regulates the differentiation of VSMCs by targeting nuclear factor- κ Bp65 (NF- κ Bp65) and the proliferative cell nuclear antigen (PCNA) [26]. Thus, it is possible that the expression change in the *MIR3142HG* gene results in the change in miR-146a level, which regulates the physiology of VSMCs, although we could not exclude a miRNA-independent role of this lncRNA. For the other lncRNAs, their roles in muscle physiology have not been reported yet. Thus, the lncRNAs identified in this analysis represent good candidates for future studies to elucidate their roles during muscle differentiation.

To identify any transcription factors with enriched binding into the promoter regions of differentially expressed lncRNAs, we used the data released by the ENCODE project, where ChIP-seq was performed for 161 major transcription factors [19]. Among the transcription factors with ChIP-seq signals, SIN3A, SMC3, and RAD21 showed significant enrichment in the promoters of lncRNAs increased in the synthetic phenotype (S2 Table). A previous report showed that the transcriptional repressor, SIN3A, is required for the development of skeletal muscle [27], although its role in smooth muscle is unknown. Further studies are required to test whether the transcriptional regulation of lncRNAs by SIN3A, and also by SMC3 and RAD21, is involved in the phenotypic switch of VSMCs.

Comparison of lncRNAs between smooth and skeletal muscles

Although smooth and skeletal muscles have different morphologies and exert dissimilar roles in the body, they share several factors during their differentiation, such as serum response factor (SRF) [28]. To identify the lncRNAs that show similar changes during the differentiation of these two muscle types, we obtained the RNA-seq data from GEO (GSE70389), which compared the transcriptome between myoblasts and myotubes during skeletal muscle differentiation [29]. We found that among 14 lncRNAs with meaningful expression in both smooth and skeletal muscles, eight lncRNAs showed a similar pattern of expression change in both muscle types (Fig 3). For example, one of the lncRNAs with the most drastic expression change during skeletal muscle differentiation, *SMILR*, has a high probability to function in skeletal muscle, although its role was only reported in smooth muscle [11]. In contrast, six lncRNAs showed opposite direction of expression change between the differentiation of smooth and skeletal muscles (Fig 3). Thus, they may exert a different effect during the differentiation of these two muscle types.

Confirmation of lncRNAs expression

To confirm the expression change of lncRNAs selected above, and to test whether the expression change of those lncRNAs is reproducible in another condition of VSMCs differentiation, we cultured HCASM cells with growth media including growth factors, and induced their differentiation with low concentration of serum and without growth factors (see Methods). Using the samples collected before and after differentiation, we confirmed the differentiation of HCASM by measuring the level of differentiation markers including CNN1 and SM22 α (Fig 4). We randomly selected five lncRNAs for the measurement, which are decreased after differentiation (Fig 2) and do not fully overlap with the nearest gene. The expression level of three lncRNAs, RP11-356I2.4, RP11-1024P17.1, and NR2F1-AS1, showed the same pattern of

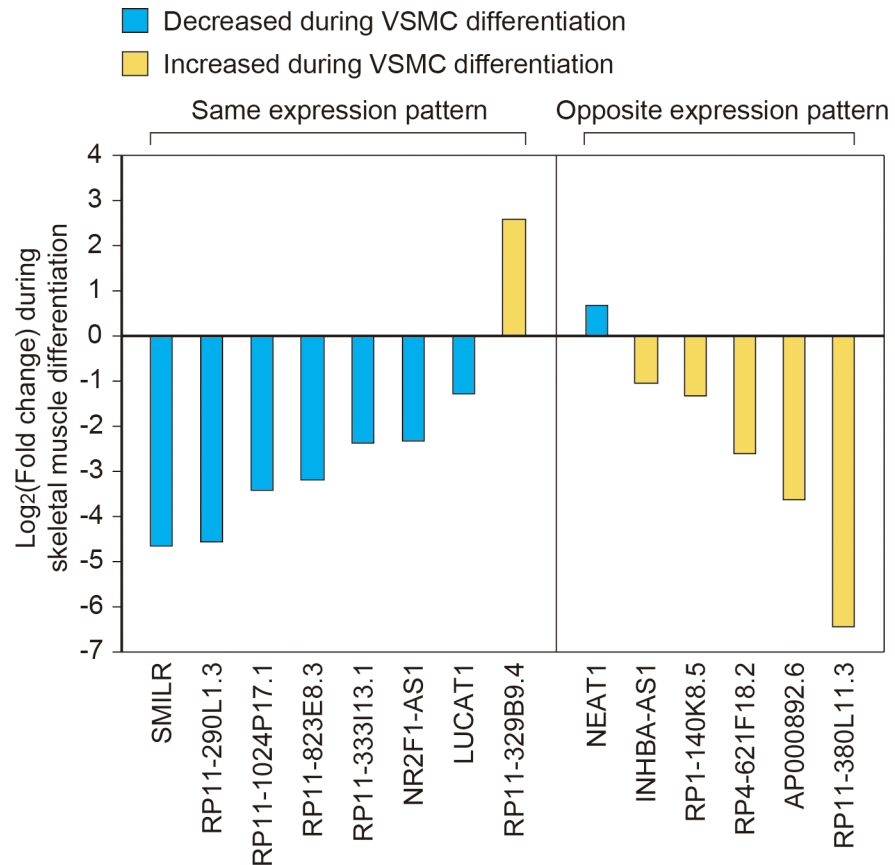


Fig 3. Comparison of long noncoding RNA (lncRNAs) expression between smooth and skeletal muscles. For those lncRNAs with reliable expression levels in both the smooth and skeletal muscle datasets, and with significant expression change during differentiation of vascular smooth muscle cells (VSMCs), their log₂(fold change) values were compared between smooth and skeletal muscle datasets. Those lncRNAs that show the same direction of expression change during differentiation in both datasets are depicted on the left. Those lncRNAs that increased during the differentiation of smooth muscle but decreased during the differentiation in skeletal muscle, or vice versa, are shown on the right.

<https://doi.org/10.1371/journal.pone.0193898.g003>

expression change also in our differentiation method (Fig 4). Thus, these lncRNAs consistently decreased in contractile or differentiation phenotype of VSMCs among diverse differentiation conditions. Two lncRNAs, LUCAT1 and RP11-333I13.1, did not decrease after differentiation in our method, suggesting that they respond differentially to diverse differentiation signals.

We utilized Coding Potential Assessment Tool (CPAT) and Coding Potential Calculator (CPC) to predict possible coding potential for three lncRNAs confirmed above [30, 31]. The result of analysis suggested that these lncRNAs have very low potential to be translated (S3 Table); although we cannot rule out the possibility that short peptide might be translated from the lncRNA sequences.

Expression correlation of lncRNAs with their neighboring genes

Previous studies that elucidated the working mechanisms of lncRNAs suggested that many lncRNAs regulate the expression of their neighboring genes in the genomic context [9]. These ‘cis-acting’ lncRNAs recruit various transcription factors or chromatin remodeling complexes to change the transcription status of nearby genes. Accordingly, the expression levels between those lncRNAs and their neighboring genes are highly correlated. To test whether the lncRNAs

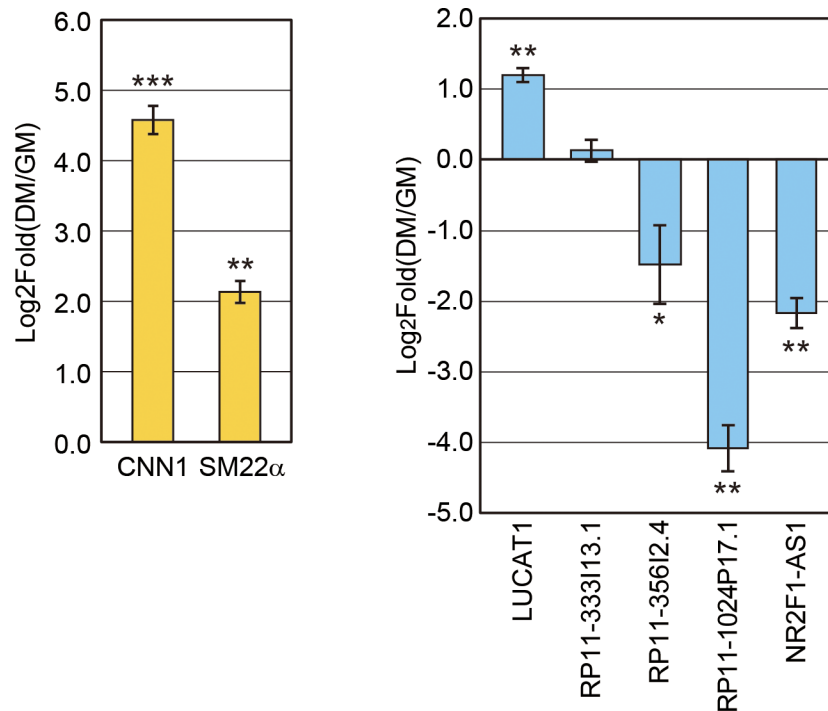


Fig 4. Expression confirmation of lncRNAs during VSMCs differentiation. The expression level of lncRNAs before and after the differentiation of VSMCs was measured. The expression of CNN1 and SM22 α , which are known to increase in contractile phenotype compared to synthetic phenotype, was used as the marker of VSMCs differentiation. GM indicates growth media while DM means differentiation media. Error bars indicate the standard error from three biological replicates (n = 3). P values were calculated from one-sided paired T test (*: $p < 0.1$, **: $p < 0.01$, ***: $p < 0.001$).

<https://doi.org/10.1371/journal.pone.0193898.g004>

identified in the present study behave in a similar way, we checked the expression of the nearest protein-coding gene for each lncRNA. We selected three lncRNAs whose expression change is confirmed in our system (Fig 4). Using the expression data from three RNA-seq datasets of smooth muscle differentiation (Fig 2) and PCR experiments (Fig 4), we calculated the correlation in fold change for lncRNA-neighboring gene pairs (Fig 5). Interestingly, all three pairs showed very high correlation in their expression changes during the differentiation of VSMCs. The data suggests that these lncRNAs may operate by regulating their nearby protein-coding genes. For example, the neighboring gene of lncRNA *RP11-1024P17.1*, is *TGFBR2* (encoding TGF beta receptor 2), which has essential roles in vascular development [32, 33]. Thus, the protein-coding genes near the identified lncRNAs tend to be functionally related to the development of VSMCs. Future studies to elucidate the connection between the lncRNA-neighboring gene pairs is essential to understand the working mechanism of these lncRNAs.

Analysis of the selected lncRNAs as competing endogenous RNAs

In addition to the role of lncRNAs as cis-acting transcriptional regulators, lncRNAs can work as competing endogenous RNAs by binding miRNAs in a sequence-specific manner and suppressing their function [9]. In this way, lncRNAs can increase the amount of target mRNAs that are originally suppressed by miRNAs. To identify miRNAs that could be regulated by selected lncRNAs above (Fig 4), we used miRcode, the database of miRNA target prediction [20]. To increase the reliability of the prediction, we utilized the expression data of miRNAs during VSMCs differentiation, which was obtained from GEO database (GSE19544) [21].

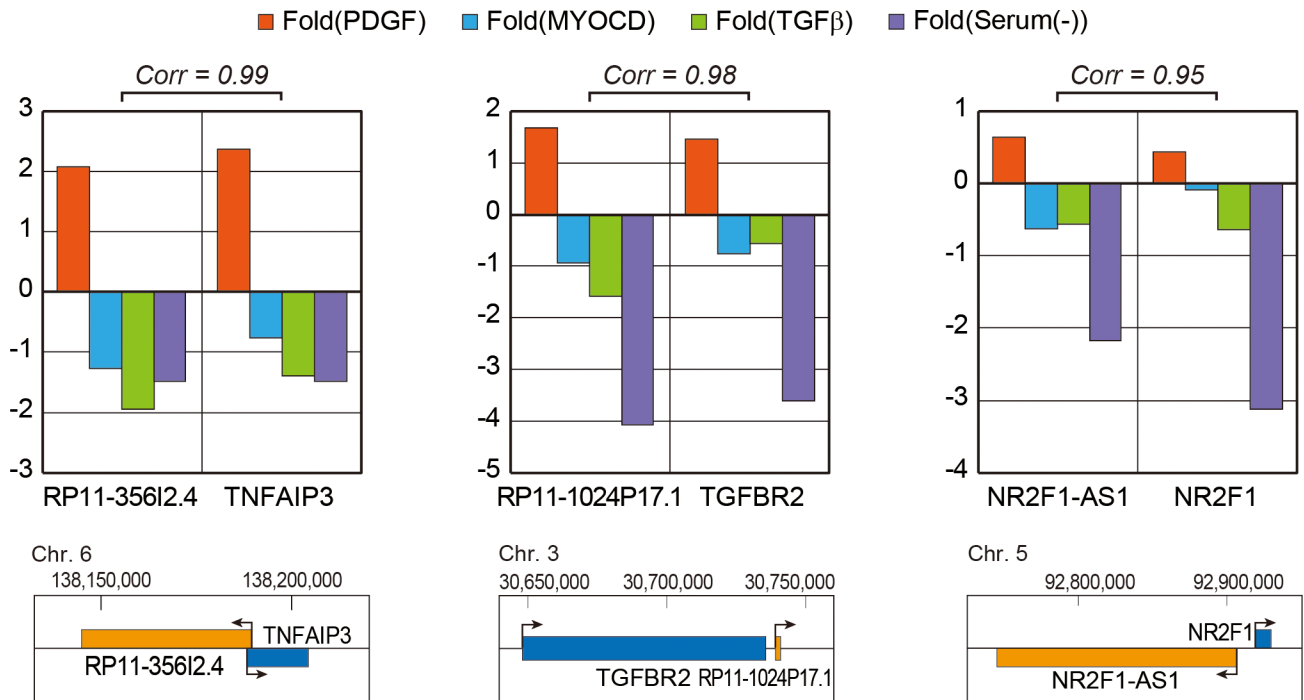


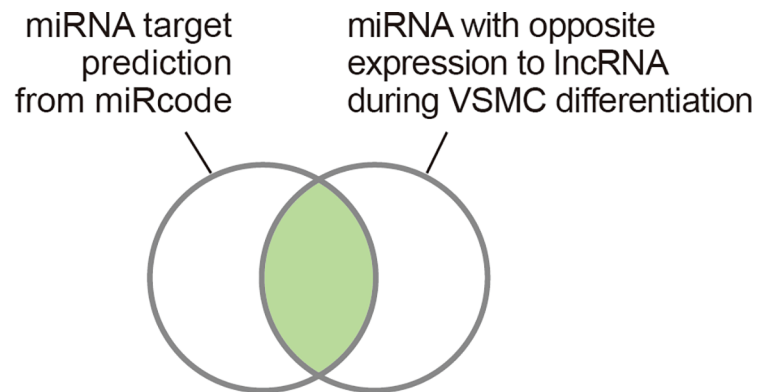
Fig 5. Expression correlation of long noncoding RNAs (lncRNAs) with their nearest genes. The log₂(fold change) of each lncRNA (left) or its neighboring protein-coding gene (right) is depicted. The fold changes for serum-starved condition were obtained by performing PCR, while other values were calculated using RNA-seq data. The correlation of fold change for each pair of lncRNA and its neighboring gene was calculated and shown above the bar graph. Below each graph, the genomic locations of the lncRNA and its neighboring gene, with their direction of transcription, are indicated. Note that the lncRNA locus is indicated by an orange box, while that of its neighboring protein-coding gene is indicated by a blue box.

<https://doi.org/10.1371/journal.pone.0193898.g005>

After obtaining the list of miRNAs that have binding sites in the lncRNA sequences, we only selected the miRNAs increased in differentiated phenotype because those three lncRNAs that we selected are decreased upon differentiation of VSMCs (Fig 4). We found several miRNAs for each lncRNA, which meet these criteria (Fig 6). Thus, these lncRNAs may regulate the function of miRNAs during phenotypic change of VSMCs by working as competing endogenous RNAs. All the miRNAs in this list were shown to be involved in the differentiation or other related phenotypes of smooth muscle cells from previous studies ([21, 34–40], suggesting the reliability of our analysis.

Finally, we searched NPInter database where the experimentally verified association between lncRNAs and proteins are documented [41]. Several interactions between lncRNAs selected above and RNA-binding proteins were found as shown in S4 Table. Among these RNA-binding proteins, ADAR1, hnRNPA1, and hnRNPA2/B1 were reported to be involved in the differentiation of smooth muscle cells [42–44]. Thus, our analyses suggest that the lncRNAs identified in this study indeed are expected to play important roles in the differentiation of VSMCs possibly through association with their binding proteins. Further researches are required to elucidate the regulatory mechanisms of these lncRNAs.

Although many lncRNAs have been discovered in diverse species and experimental models, their roles and pathological effects in muscles have not been studied as extensively as those related to other diseases, such as cancer. One of the issues in discovering lncRNAs or other types of noncoding RNAs in an experimental model is their inconsistency when analyzed using different approaches. Thus, lncRNAs identified in one approach may not be replicated in another approach, although both approaches are aimed at the same experimental model.



RP11-1024P17.1

miRNA	# sites	Log ₂ Fold(48hr)	Log ₂ Fold(72hr)
miR-26	1	1.882	0.893
miR-30-5p	1	1.104	0.605
miR-143	1	2.065	1.425

RP11-356I2.4

miRNA	# sites	Log ₂ Fold(48hr)	Log ₂ Fold(72hr)
miR-23	1	0.936	0.533
miR-204	2	0.990	0.727

NR2F1-AS1

miRNA	# sites	Log ₂ Fold(48hr)	Log ₂ Fold(72hr)
let-7/miR-98	1	1.056	0.574
miR-34-5p	1	1.747	1.016
miR-143	1	2.065	1.425
miR-199-5p	2	1.329	0.940

Fig 6. Analysis of lncRNAs as miRNA regulators. We intersected the list of lncRNA-targeting miRNAs predicted from miRcode with the list of miRNAs that shows opposite expression patterns to lncRNAs during VSMC differentiation, as depicted in the diagram (see [Methods](#)). For each lncRNA, selected miRNAs with binding sites at corresponding lncRNA and with inverse correlation in expression to the same lncRNA are shown. The number of miRNA-binding sites in each lncRNA sequence is indicated. The expression change of miRNAs between before and after the differentiation of VSMCs was re-calculated using the miRNA microarray data from GSE19544. The fold ratios from two different time points are indicated.

<https://doi.org/10.1371/journal.pone.0193898.g006>

The comparison of the expression profile of lncRNAs in the differentiation model of VSMCs confirmed this issue ([Fig 1C](#)). Although all the different treatments, including PDGF,

MYOCD, and TGF β , mimic the differentiation conditions of VSMCs, most lncRNAs showed inconsistent changes among these conditions. In addition, two of five lncRNAs that were consistently changed from these conditions showed different expression pattern in our serum withdrawal condition. In our previous study to identify miRNAs important in the progression of Alzheimer's disease, we improved this problem by integrating various RNA-seq data sets from human patients, a mouse model, and a cellular model, and identified commonly changed miRNAs in these models [45]. In the present study, we applied a similar approach to identify lncRNAs that may have important roles in VSMCs differentiation by integrating the RNA-seq data from various differentiation conditions for VSMCs and confirming their expression change in yet another condition. From this analysis, we discovered dozens of lncRNAs with possible involvement in the phenotypic change of VSMCs, most of which have no previous reports of a link to the physiology of VSMCs (Fig 2). Moreover, for three lncRNAs, we confirmed the expression change in our differentiation system and analyzed their possible regulatory mechanism. Therefore, the list of lncRNAs identified in this study will be a useful resource for future research into the gene regulatory network during muscle differentiation.

Supporting information

S1 Fig. Heat map of long noncoding RNAs (lncRNAs) whose expression was increased in the contractile phenotype. The cluster indicated in Fig 1 with yellow box is magnified.
(PDF)

S2 Fig. Heat map of long noncoding RNAs (lncRNAs) whose expression is increased in the synthetic phenotype. The cluster indicated in Fig 1 by a blue box is magnified.
(PDF)

S1 Table. Long noncoding RNAs (lncRNAs) with differential expression during phenotypic change of vascular smooth muscle cells (VSMCs). For each lncRNA, their genomic locus (hg19), the fragments per kilobase of exon model per million mapped fragments (FPKM) values in each sample, log₂ (fold change), and the *p* values from two-sided T test for each dataset, are shown. For those lncRNAs whose promoters were analyzed, the position of the transcription start sites and the strand of transcription are included. For the lncRNAs with incomplete annotation (red letters), transcription start sites were identified by manual curation using the information from RefSeq annotation. Only the significant lncRNAs from Fig 2 were included.
(XLSX)

S2 Table. Long noncoding RNAs (lncRNAs) with chromatin immunoprecipitation sequencing (ChIP-seq) signals of transcription factors in their promoter regions. Among the transcription factors included in the ChIP-seq data from ENCODE, we identified those that were significantly enriched in the promoters of lncRNAs that increased in the synthetic phenotype compared with those increased in the contractile phenotype (one-sided Fisher's exact test). This analysis resulted in three transcription factors with *p* values lower than 0.1. There was no transcription factor enriched in the lncRNA promoters of the contractile phenotype against those of the synthetic phenotype.
(XLSX)

S3 Table. Bioinformatics analysis of coding potential of selected lncRNAs. The coding potential of selected lncRNAs in Figs 5 and 6 was analyzed at (A) CPAT and (B) CPC web servers [30, 31].
(XLSX)

S4 Table. Interaction of selected lncRNAs with proteins. We used NPInter database to search the experimentally verified association between lncRNAs and proteins [41], and listed the pairs of lncRNA and interacting protein. The experimental method to identify these associations was also shown.

(XLSX)

S5 Table. Primer sequences used in this study.

(XLSX)

Acknowledgments

We thank the members of the Department of Biochemistry and the Department of Pharmacology in Chonnam National University Medical School for their helpful discussions.

Author Contributions

Conceptualization: Young-Kook Kim.

Data curation: Young-Kook Kim.

Formal analysis: Yeong-Hwan Lim, Jaetaek Kim, Woo Jin Park, Hyun Kook, Young-Kook Kim.

Funding acquisition: Jaetaek Kim, Woo Jin Park, Hyun Kook, Young-Kook Kim.

Methodology: Duk-Hwa Kwon.

Supervision: Young-Kook Kim.

Writing – original draft: Young-Kook Kim.

Writing – review & editing: Yeong-Hwan Lim, Jaetaek Kim, Woo Jin Park, Hyun Kook, Young-Kook Kim.

References

1. Rensen SS, Doevendans PA, van Eys GJ. Regulation and characteristics of vascular smooth muscle cell phenotypic diversity. *Neth Heart J*. 2007; 15(3):100–8. PMID: [17612668](#)
2. Davis-Dusenbery BN, Wu C, Hata A. Micromanaging vascular smooth muscle cell differentiation and phenotypic modulation. *Arterioscler Thromb Vasc Biol*. 2011; 31(11):2370–7. <https://doi.org/10.1161/ATVBAHA.111.226670> PMID: [22011749](#)
3. Owens GK, Kumar MS, Wamhoff BR. Molecular regulation of vascular smooth muscle cell differentiation in development and disease. *Physiol Rev*. 2004; 84(3):767–801. <https://doi.org/10.1152/physrev.00041.2003> PMID: [15269336](#)
4. Orr AW, Hastings NE, Blackman BR, Wamhoff BR. Complex regulation and function of the inflammatory smooth muscle cell phenotype in atherosclerosis. *J Vasc Res*. 2010; 47(2):168–80. <https://doi.org/10.1159/000250095> PMID: [19851078](#)
5. Chal J, Pourquie O. Making muscle: skeletal myogenesis in vivo and in vitro. *Development*. 2017; 144(12):2104–22. <https://doi.org/10.1242/dev.151035> PMID: [28634270](#)
6. Tedesco FS, Dellavalle A, Diaz-Manera J, Messina G, Cossu G. Repairing skeletal muscle: regenerative potential of skeletal muscle stem cells. *J Clin Invest*. 2010; 120(1):11–9. <https://doi.org/10.1172/JCI40373> PMID: [20051632](#)
7. Kim YK. Extracellular microRNAs as Biomarkers in Human Disease. *Chonnam Med J*. 2015; 51(2):51–7. <https://doi.org/10.4068/cmj.2015.51.2.51> PMID: [26306299](#)
8. Coll-Bonfill N, de la Cruz-Thea B, Pisano MV, Musri MM. Noncoding RNAs in smooth muscle cell homeostasis: implications in phenotypic switch and vascular disorders. *Pflugers Arch*. 2016; 468(6):1071–87. <https://doi.org/10.1007/s00424-016-1821-x> PMID: [27109570](#)
9. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet*. 2014; 15(1):7–21. <https://doi.org/10.1038/nrg3606> PMID: [24296535](#)

10. Bell RD, Long X, Lin M, Bergmann JH, Nanda V, Cowan SL, et al. Identification and initial functional characterization of a human vascular cell-enriched long noncoding RNA. *Arterioscler Thromb Vasc Biol.* 2014; 34(6):1249–59. <https://doi.org/10.1161/ATVBAHA.114.303240> PMID: 24578380
11. Ballantyne MD, Pinel K, Dakin R, Vesey AT, Diver L, Mackenzie R, et al. Smooth Muscle Enriched Long Noncoding RNA (SMILR) Regulates Cell Proliferation. *Circulation.* 2016; 133(21):2050–65. <https://doi.org/10.1161/CIRCULATIONAHA.115.021019> PMID: 27052414
12. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 2013; 41(Database issue):D991–5. <https://doi.org/10.1093/nar/gks1193> PMID: 23193258
13. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014; 30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
14. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013; 29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635> PMID: 23104886
15. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009; 25(16):2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
16. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* 2012; 22(9):1760–74. <https://doi.org/10.1101/gr.135350.111> PMID: 22955987
17. Tyner C, Barber GP, Casper J, Clawson H, Diekhans M, Eisenhart C, et al. The UCSC Genome Browser database: 2017 update. *Nucleic Acids Res.* 2017; 45(D1):D626–D34. <https://doi.org/10.1093/nar/gkw1134> PMID: 27899642
18. Jeong G, Lim YH, Kim YK. Precise mapping of the transcription start sites of human microRNAs using DROSHA knockout cells. *BMC Genomics.* 2016; 17(1):908. <https://doi.org/10.1186/s12864-016-3252-7> PMID: 27835943
19. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012; 489(7414):57–74. <https://doi.org/10.1038/nature11247> PMID: 22955616
20. Jeggari A, Marks DS, Larsson E. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics.* 2012; 28(15):2062–3. <https://doi.org/10.1093/bioinformatics/bts344> PMID: 22718787
21. Leeper NJ, Raiesdana A, Kojima Y, Chun HJ, Azuma J, Maegdefessel L, et al. MicroRNA-26a is a novel regulator of vascular smooth muscle cell function. *J Cell Physiol.* 2011; 226(4):1035–43. <https://doi.org/10.1002/jcp.22422> PMID: 20857419
22. Zhao J, Wu W, Zhang W, Lu YW, Tou E, Ye J, et al. Selective expression of TSPAN2 in vascular smooth muscle is independently regulated by TGF-beta1/SMAD and myocardin/serum response factor. *FASEB J.* 2017; 31(6):2576–91. <https://doi.org/10.1096/fj.201601021R> PMID: 28258189
23. Ounzain S, Micheletti R, Arnan C, Plaisance I, Cecchi D, Schroen B, et al. CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis. *J Mol Cell Cardiol.* 2015; 89(Pt A):98–112. <https://doi.org/10.1016/j.yjmcc.2015.09.016> PMID: 26423156
24. Brock M, Schuoler C, Leuenberger C, Buhlmann C, Haider TJ, Vogel J, et al. Analysis of hypoxia-induced noncoding RNAs reveals metastasis-associated lung adenocarcinoma transcript 1 as an important regulator of vascular smooth muscle cell proliferation. *Exp Biol Med (Maywood).* 2017; 242(5):487–96.
25. Zhou L, Sun K, Zhao Y, Zhang S, Wang X, Li Y, et al. Linc-YY1 promotes myogenic differentiation and muscle regeneration through an interaction with the transcription factor YY1. *Nat Commun.* 2015; 6:10026. <https://doi.org/10.1038/ncomms10026> PMID: 26658965
26. Dong S, Xiong W, Yuan J, Li J, Liu J, Xu X. MiRNA-146a regulates the maturation and differentiation of vascular smooth muscle cells by targeting NF-kappaB expression. *Mol Med Rep.* 2013; 8(2):407–12. <https://doi.org/10.3892/mmr.2013.1538> PMID: 23784108
27. van Oevelen C, Bowman C, Pellegrino J, Asp P, Cheng J, Parisi F, et al. The mammalian Sin3 proteins are required for muscle development and sarcomere specification. *Mol Cell Biol.* 2010; 30(24):5686–97. <https://doi.org/10.1128/MCB.00975-10> PMID: 20956564
28. Coletti D, Daou N, Hassani M, Li Z, Parlakian A. Serum Response Factor in Muscle Tissues: From Development to Ageing. *Eur J Transl Myol.* 2016; 26(2):6008. <https://doi.org/10.4081/ejtm.2016.6008> PMID: 27478561

29. Martone J, Briganti F, Legnini I, Morlando M, Picillo E, Sthandier O, et al. The lack of the Celf2a splicing factor converts a Duchenne genotype into a Becker phenotype. *Nat Commun*. 2016; 7:10488. <https://doi.org/10.1038/ncomms10488> PMID: 26796035
30. Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res*. 2013; 41(6):e74. <https://doi.org/10.1093/nar/gkt006> PMID: 23335781
31. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res*. 2007; 35(Web Server issue):W345–9. <https://doi.org/10.1093/nar/gkm391> PMID: 17631615
32. Hellbach N, Weise SC, Vezzali R, Wahane SD, Heidrich S, Roidl D, et al. Neural deletion of Tgfr2 impairs angiogenesis through an altered secretome. *Hum Mol Genet*. 2014; 23(23):6177–90. <https://doi.org/10.1093/hmg/ddu338> PMID: 24990151
33. Li W, Li Q, Jiao Y, Qin L, Ali R, Zhou J, et al. Tgfr2 disruption in postnatal smooth muscle impairs aortic wall homeostasis. *J Clin Invest*. 2014; 124(2):755–67. <https://doi.org/10.1172/JCI69942> PMID: 24401272
34. Hashemi Gheinani A, Burkhard FC, Rehrauer H, Aquino Fournier C, Monastyrskaya K. MicroRNA MiR-199a-5p regulates smooth muscle cell proliferation and morphology by targeting WNT2 signaling pathway. *J Biol Chem*. 2015; 290(11):7067–86. <https://doi.org/10.1074/jbc.M114.618694> PMID: 25596533
35. Wang TM, Chen KC, Hsu PY, Lin HF, Wang YS, Chen CY, et al. microRNA let-7g suppresses PDGF-induced conversion of vascular smooth muscle cell into the synthetic phenotype. *J Cell Mol Med*. 2017; 21(12):3592–601. <https://doi.org/10.1111/jcmm.13269> PMID: 28699690
36. Zong Y, Wu P, Nai C, Luo Y, Hu F, Gao W, et al. Effect of MicroRNA-30e on the Behavior of Vascular Smooth Muscle Cells via Targeting Ubiquitin-Conjugating Enzyme E21. *Circ J*. 2017; 81(4):567–76. <https://doi.org/10.1253/circj.CJ-16-0751> PMID: 28123167
37. Yu X, Zhang L, Wen G, Zhao H, Luong LA, Chen Q, et al. Upregulated sirtuin 1 by miRNA-34a is required for smooth muscle cell differentiation from pluripotent stem cells. *Cell Death Differ*. 2015; 22(7):1170–80. <https://doi.org/10.1038/cdd.2014.206> PMID: 25526086
38. Cui RR, Li SJ, Liu LJ, Yi L, Liang QH, Zhu X, et al. MicroRNA-204 regulates vascular smooth muscle cell calcification in vitro and in vivo. *Cardiovasc Res*. 2012; 96(2):320–9. <https://doi.org/10.1093/cvr/cvs258> PMID: 22871591
39. Cordes KR, Sheehy NT, White MP, Berry EC, Morton SU, Muth AN, et al. miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature*. 2009; 460(7256):705–10. <https://doi.org/10.1038/nature08195> PMID: 19578358
40. Iaconetti C, De Rosa S, Polimeni A, Sorrentino S, Gareri C, Carino A, et al. Down-regulation of miR-23b induces phenotypic switching of vascular smooth muscle cells in vitro and in vivo. *Cardiovasc Res*. 2015; 107(4):522–33. <https://doi.org/10.1093/cvr/cvv141> PMID: 25994172
41. Hao Y, Wu W, Li H, Yuan J, Luo J, Zhao Y, et al. NPInter v3.0: an upgraded database of noncoding RNA-associated interactions. *Database (Oxford)*. 2016; 2016.
42. Fei J, Cui XB, Wang JN, Dong K, Chen SY. ADAR1-Mediated RNA Editing, A Novel Mechanism Controlling Phenotypic Modulation of Vascular Smooth Muscle Cells. *Circ Res*. 2016; 119(3):463–9. <https://doi.org/10.1161/CIRCRESAHA.116.309003> PMID: 27199464
43. Wang G, Xiao Q, Luo Z, Ye S, Xu Q. Functional impact of heterogeneous nuclear ribonucleoprotein A2/B1 in smooth muscle differentiation from stem cells and embryonic arteriogenesis. *J Biol Chem*. 2012; 287(4):2896–906. <https://doi.org/10.1074/jbc.M111.297028> PMID: 22144681
44. Huang Y, Lin L, Yu X, Wen G, Pu X, Zhao H, et al. Functional involvements of heterogeneous nuclear ribonucleoprotein A1 in smooth muscle differentiation from stem cells in vitro and in vivo. *Stem Cells*. 2013; 31(5):906–17. <https://doi.org/10.1002/stem.1324> PMID: 23335105
45. Song J, Kim YK. Identification of the Role of miR-142-5p in Alzheimer's Disease by Comparative Bioinformatics and Cellular Analysis. *Front Mol Neurosci*. 2017; 10:227. <https://doi.org/10.3389/fnmol.2017.00227> PMID: 28769761