The LINC00961 transcript and its encoded micropeptide SPAAR regulate endothelial cell function

Citation for published version:

Digital Object Identifier (DOI):
10.1093/cvr/cvaa008

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Cardiovascular Research
The LINC00961 transcript and its encoded micropeptide SPAAR regulate endothelial cell function

Short title: Angiogenic roles of the LINC00961/SPAAR locus.

Helen L. Spencer 1*, Rachel Sanders 1*, Mounia Boulberdaa 1, Marco Meloni 1, Amy Cochrane 1, Ana-Mishel Spiroski 1, Joanne Mountford 3, Costanza Emanueli 2, Andrea Caporali 1, Mairi Brittan 1, Julie Rodor 1, Andrew H. Baker 1,3#. 

* Authors have contributed equally.

1 University/BHF Centre for Cardiovascular Science, Queens Medical Research Institute, University of Edinburgh, EH16 4TJ, UK.

2 National Heart and Lung Institute, Vascular Sciences and 5 Cardiac Function, Imperial Centre for Translational and Experimental Medicine, Imperial College London, London, W12 0NN, UK.

3 Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK.

# Correspondence: Andrew H Baker, University of Edinburgh, Centre for Cardiovascular Sciences, Queen’s Medical Research Institute, 47 Little France Crescent, Edinburgh UK, EH16 4TJ, Tel: 0131 24 26728, Email: Andy.Baker@ed.ac.uk.

Manuscript Category: Original Article

Total word count: 7915

Author’s contribution:
A.H.B conceived the study. A.H.B, H.L.S, R.S, M.B and J.R designed experiments and interpreted data. H.L.S, R.S, M.B, M.M. C.R.P performed experiments. J.R performed the bioinformatics analysis, A.C, M.B, J.M, J.R and A.H.B supervised the research. A.H.B, H.L.S, R.S and J.R wrote the manuscript. All the authors discussed the data and edited the manuscript.

Statements:
The manuscript is not submitted elsewhere or under consideration for publication. All authors have read and agreed with the submission of the manuscript.

© The Author(s) 2020. Published by Oxford University Press on behalf of the European Society of Cardiology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Abstract

Aims: Long non-coding RNAs (lncRNAs) play functional roles in physiology and disease, yet understanding of their contribution to endothelial cell (EC) function is incomplete. We identified lncRNAs regulated during EC differentiation and investigated the role of LINC00961 and its encoded micropeptide SPAAR in EC function.

Methods and Results: Deep sequencing of human embryonic stem cell differentiation to ECs was combined with ENCODE RNA-seq data from vascular cells, identifying 278 endothelial enriched genes, including 6 lncRNAs. Expression of LINC00961, first annotated as a lncRNA but reassigned as a protein coding gene for the SPAAR micropeptide, was increased during the differentiation and was EC-enriched. LINC00961 transcript depletion significantly reduced EC adhesion, tube formation, migration, proliferation, and barrier integrity in primary ECs. Overexpression of the SPAAR open reading frame increased tubule formation, however overexpression of the full length transcript did not, despite production of SPAAR. Further, overexpression of an ATG mutant of the full length transcript reduced network formation, suggesting a bona fide non-coding RNA function of the transcript with opposing effects to SPAAR. As the LINC00961 locus is conserved in mouse, we generated a LINC00961 locus knockout (KO) mouse that underwent hind limb ischaemia to investigate the angiogenic role of this locus in vivo. In agreement with in vitro data, KO animals had a reduced capillary density in the ischaemic adductor muscle after 7 days. Finally, to characterise LINC00961 and SPAAR independent functions in ECs, we performed pull-downs of both molecules and identified protein binding partners. LINC00961 RNA binds the G-actin sequestering protein thymosin beta-4x (Tβ4) and Tβ4 depletion phenocopied the overexpression of the ATG mutant. SPAAR binding partners included the actin binding protein, SYNE1.

Conclusion: The LINC00961 locus regulates EC function in vitro and in vivo. The gene produces two molecules with opposing effects on angiogenesis: SPAAR and LINC00961.

Words 298

Translational Perspective (100 words):
Treatment of ischemic conditions remains a major cardiovascular health burden. Identification of genes and non-coding RNAs that regulate the function of the vascular endothelium is important to understand and evolve potential new strategies that might enhance vascular regeneration. Here, we describe and dissect the functional importance of a micropeptide-encoding RNA transcript in the vascular endothelium, and demonstrate that both the RNA and the peptide regulate endothelial biology. Modulation of this axis may be a novel approach to regulate angiogenesis.

Non-standard Abbreviations and Acronyms
aa: Amino acids
α-sm: α-smooth muscle actin
bp: Base pairs
DsiRNA: Dicer substrate siRNA
EC: Endothelial cell
ECIS: Electric cell-substrate impedance sensing assay
ENCODE: Encyclopedia of DNA Elements
FISH: Fluorescent in situ hybridisation
GO: Gene Ontology
hESC: Human embryonic stem cell
HSVEC: Human saphenous vein endothelial cell
HUVEC: Human umbilical vein endothelial cell
IB4: Isolectin B4
KD: Knock down
KO: Knock out
**lncRNA: Intergenic lncRNA**

**IncRNA: Long non-coding RNA**

**LV: Lentiviral**

**MP: Day 3 mesodermal population**

**ncRNA: non-coding RNA**

**ORF: Open reading frame**

**PCA: Principal component analysis**

**SMC: Smooth muscle cell**

**SPAAR: Small regulatory polypeptide of amino acid response**

**TP4: Thymosin beta-4x**

**UBC: Ubiquitin C**

**WT: Wild type**

(iii) **INTRODUCTION**

The endothelium is a heterogeneous organ system that regulates homeostasis of the vasculature and represents a permeable monolayer barrier between the vessel wall and the blood. Endothelial cells (ECs) regulate and adapt to shear stress, leukocyte extravasation, blood clotting, inflammation, vascular tone, extracellular matrix deposition, vasoconstriction/vasodilation, and angiogenesis. During angiogenesis, ECs become activated and undergo sprouting, proliferation, migration along a gradient of pro-angiogenic factors (eg. VEGF, FGF, PDGF), and Anastomose to form new capillaries before returning to their quiescent state. Aberrant activation however, leads to EC dysfunction that can cause systemic vascular pathology. This uncontrolled activation is a significant factor contributing to coronary artery disease, diabetes, hypertension patients, hypercholesterolemia, lupus, and has been reported as increased in smokers.

Several groups have demonstrated the ability to differentiate ECs from human embryonic stem cells (hESC). This protocol yields ECs that are relatively immature and express genes that are somewhat distinct from those of mature ECs from various vascular beds, highlighting the importance of understanding the molecular mechanisms controlling both general and specialised EC differentiation, specification, and function. These derived ECs have been extensively proven to be functional both in vitro, by the ability to form capillary like networks on Matrigel and in vivo, by their ability to improve vascular density and perfusion in a murine model of hind limb ischaemia (HLI). These data provide evidence of the benefits to hESC-derived EC for therapeutics and as a model to characterise early vascular development.

Data from the human Encyclopedia of DNA Elements (ENCODE) project indicates that approximately 93% of the genome is transcribed, with less than 2% encoding protein sequences. Currently, these non-coding RNAs (ncRNAs) are classified based on size, into long non-coding RNAs (lncRNAs) >200bp and small ncRNAs < 200bps. LncRNAs correspond to a heterogeneous class of genes, with subtypes classified based on neighbouring protein-coding genes. In particular, lncRNAs are intergenic lncRNAs with no overlap with protein-coding genes. While some lncRNAs regulate in cis their protein-coding neighbours expression, a large range of trans-functions have been reported including chromatin remodelling, transcriptional and post-transcriptional regulation, translation control and regulation of protein activity. LncRNAs show spatio-temporal expression, and are poorly conserved between species, however, to date only a few of the lncRNAs known to exist have been functionally characterised. Recent literature highlights the important functions of lncRNAs as regulators of the cardiovascular system. In the vascular endothelium, TIE1-AS1 was the first described endothelial specific IncRNA, involved in modulating TIE-1 expression and regulating endothelial vessel formation. A comprehensive transcriptome analysis of early cardiovascular development revealed the regulation of several lncRNAs and led to the characterisation of ALIEN and PUNISHER. Recently, the hypoxia-induced lncRNA, GATA6-AS, was shown to epigenetically regulate angiogenesis through its interaction with the epigenetic regulator LOXL2.
The “non-coding” property of some lncRNAs has been disputed by the discovery of small open reading frames (ORFs) in some lncRNA transcripts, able to generate functional micropeptides. For example, LINC00948 has been reclassified as a protein-coding gene, as it encodes myoregulin, which inhibits the calcium ATPase SERCA in muscle. Similarly, the micropeptide DWORF encoded by lncRNA NONMMUG206737 activates the SERCA pump. Noteworthy to this study, a conserved micropeptide termed small regulatory polyepitope of amino acid response (SPAAR) was recently shown to be encoded by the LINC00961 locus. SPAAR attenuates lysosomal v-ATPases interaction with mTORC1 under amino acid stimulation and modulates skeletal muscle regeneration following cardiotoxin injury. These studies focused on the function of the derived micropeptide, however, some micropeptides have been shown to be expressed from lncRNAs with previously characterised non-coding functions, suggesting the possibility of bi-functional loci.

We identified the LINC00961/SPAAR locus as EC enriched, and sought to identify the role of this micropeptide-encoding gene. This led to dissection of the contribution of the LINC00961 RNA transcript itself and the SPAAR micropeptide on endothelial function. LINC00961 RNA was found to act as a bona fide lncRNA that inhibited angiogenesis and bound to the known angiogenic and actin binding protein thymosin beta 4 (Tß4). Whereas SPAAR was found to be pro-angiogenic and bound to another actin binding protein, SYNE1.

(iv) METHODS

Endothelial cell isolation and cell culture

All donated tissues have been obtained under proper informed consent and the investigation conforms with principles in the Declaration of Helsinki. Human saphenous vein endothelial cells (HSVECs) were obtained by enzymatic collagenase digestion of human saphenous veins (Ethics 15/ES/0094). Human umbilical vein endothelial cells (HUVECs) were obtained from Lonza (Basel Switzerland).

RNA-Seq of human embryonic stem cell differentiation to endothelial cells

A previously published protocol was employed to generate endothelial cells from H9 human embryonic stem cells. RNA-Seq analysis was performed as previously described with minor modifications. Ensembl GRCh38 was used for transcriptome annotation. Read counts for each gene were obtained using HTSeq. The differential expression was analysed using DESeq2. RNA-seq data are deposited at the Gene expression Omnibus as GSE118106.

Expression data from several human endothelial and smooth muscle cell (SMC) lines was obtained from the Encyclopedia of DNA Element (ENCODE) Consortium. The list of analysed data and their abbreviated name can be found in Supplementary Table 1. Candidate filtering was done as follow: (a) Genes enriched in day7 EC versus hESC and non-EC day sample based on a LogFC>1, padj<0.01, FPKM>2 (b) Genes up-regulated in HSVEC versus hESC (LogFC>1, padj<0.01, FPKM>2) (c) Genes expressed in ENCODE ECs (min of 2 FPKM in 10 samples) (d) Enriched expressed in ENCODE ECs versus ENCODE SMCs (2 fold enrichment between the average expression in ECs and SMCs).

HUVEC transfection and phenotype analysis

All phenotypes were assessed in Human umbilical vein endothelial cells (HUVECs) at 24 hours after transfection with dicer substrate siRNA (dsiRNA) or infection with lentiviral constructs (details of reagents and protocols in Supplementary Methods). In vitro tubule network formation was assessed using Matrigel (Corning, USA) according to the manufacturer’s protocol. Proliferation was assessed using the Click-it EdU 488 Proliferation assay (Life Technologies, UK). Migration and endothelial barrier function assays were performed using an Electric Cell-substrate Impedance Sensing (ECIS) machine (Applied BioPhysics) and cell viability assessed with a FTTC Annexin V Detection Kit with PI (BioLegend).
Hind limb ischaemia
All animal experiments were performed in accordance with the Animals (Scientific Procedures) Act (UK) 1986 and under the auspices of UK Home Office Project and Personal Licenses held within The University of Edinburgh facilities. LINC00961Δ mouse line was obtained from Taconic®. Validation of genotype was two-fold. Ear clip samples from pups were sent to Transnetyx® for genotyping, and in house validation was also carried out using qRT-PCR on mRNA extracted from the kidney. Surgical procedures were performed under inhaled general anaesthesia (isoflurane at 5% for induction and 1-2% for maintenance) and with appropriate peri-operative analgesic cover (subcutaneous injection of buprenorphine at 0.05mg/kg). Unilateral hind limb ischaemia (HLI) was surgically induced by left femoral artery ligation at two points and cauterisation of this segment of artery, leaving the femoral vein and nerve untouched. Mice were maintained for 7 days after surgery. Male LINC00961Δ and wild type (WT) littermates on the C57Bl/b6NTAC were studied at 11 weeks of age. Animals were euthanised with pentobarbital (160mg/kg) given by intraperitoneal injection. Tissues were perfusion fixed with PBS at 6ml/min with a micro pump and then with 4% paraformaldehyde at 6ml/min.

Pull-down
LINC00961 RNA pull-down was carried out with 50 pmol biotinylated IncRNA, obtained using the T7 RiboMAX Express Large Scale RNA Production System (Promega, UK). The biotinylated IncRNA was incubated with streptavidin magnetic beads and 20 μg of HUVECs protein lysate, using the Pierce Mag RNA Protein Pull-down kit (Thermo Scientific). For the SPAAR pull-down HUVECs expressing either LV-Null, LV-SPAAR untagged or LV-SPAAR- HA tagged were cultured in EGM-2 media. Immunoprecipitation with either anti-IgG or anti- HA antibody was performed in two replicates. SPAAR binding partners were defined as proteins detected in the two pull-down replicates and with a 2-fold enrichment compared to the IgG pull-down controls or pull-down in cells not overexpressing HA-tagged SPAAR. Keratin contaminants and unknown proteins were removed from the final candidate list.

Statistical analysis
Statistical analysis was performed as described in the figure legends using GraphPad Prism version 5.0. Data are expressed as mean ± SEM. Comparisons between 2 groups were analysed using 2-tailed unpaired Student’s t-tests. Comparisons between more than 2 groups were analysed using One-way ANOVA. For qRT-PCR analysis, graphs display the expression relative to the housekeeping gene based on the double ΔCt analysis while the statistical analyses were done on ΔCt values. For data represented as fold change, the statistical analysis was done on the Log2 Fold Change using a One Sample t-test.

RESULTS

Identification of endothelial cell enriched genes
To identify genes specifically induced during endothelial fate specification and differentiation, we utilised an embryoid body-based protocol to generate ECs from human embryonic stem cells (hESCs) (Figure 1A) 25. This protocol was previously shown to generate functional hESC-EC, expressing CD144 and CD31 and able to form tube-like structures on Matrigel 7. RNA-seq was performed (45 million paired end reads per sample) on ribosomal RNA depleted libraries from several replicates of the different cell populations (Figure 1A). Principal component analysis (PCA) demonstrated tight clustering of replicates and segregation of populations (Figure 1B). The purified EC samples obtained at day 7 (d7 EC) were closer to the human saphenous vein EC (HSVEC) samples in the PCA plot, but clearly clustered separately suggesting the immaturity of this EC population (Figure 1B). As expected, hESC pluripotency markers showed a down-regulation after day 3 of differentiation while mesoderm markers are up-regulated. We confirmed the expression of several endothelial markers in the d7 EC population but also showed the expression of arterial, venous and lymphatic phenotype markers, suggesting endothelial heterogeneity (Supplementary Figure 1A). As expected, we observed a high overlap between the genes up-regulated in d7 EC versus hESC and the genes up-regulated in HSVEC versus hESC (Supplementary Figure 1B), validating their endothelial identity.
To identify genes important for endothelial identity and function, we focused on candidates showing high expression in immature and mature ECs. We specifically selected 409 genes enriched in the day7 EC population but also expressed in our HSVEC samples. Then, we took advantage of RNA-seq data from the ENCODE consortium to assess their expression in several EC lines from different origins but also in SMCs. We retrieved a list of 278 genes with high expression in ECs and lower expression in SMCs (Supplementary Table 2). This list contains known markers of ECs including PECAM1, CDH5, and ERG, and the Gene Ontology (GO) analysis revealed the enrichment of terms related to vessel development and angiogenesis (Supplementary Figure 1C).

**LINC00961 is enriched in immature and mature ECs**

Among the 278 genes enriched in immature and mature ECs, we found 6 lncRNAs: 3 antisense lncRNAs and 3 intergenic lncRNAs (Figure 2A and B). While antisense RNAs often regulate the expression of their sense genes, intergenic lncRNAs have function generally unrelated to their neighbouring protein-coding genes. From the 3 intergenic lincRNA, LINC00961 is the only one conserved in mouse (Figure 2C). LINC00961 is located on chromosome 9 and while LINC00961 transcript expression was detected in the d7 EC population and HSVECs with a read profile confirming a two-exon gene structure, neighbouring HRCT1 expression was restricted to HSVECs (Figure 2C). Although LINC00961 was initially annotated as a lncRNA, the locus encodes a small ORF in the second exon and has been re-annotated as a protein coding gene. Interestingly, the peptide was independently identified based on a proteomic strategy and termed SPAAR for small regulatory polypeptide of amino acid response. To validate the RNA-seq, LINC00961 gene expression was evaluated by qRT-PCR in the same sample set used for RNA-seq, which demonstrated the same profile of expression (Supplementary Figure 2).

**LINC00961/SPAAR gene silencing affects endothelial function**

To assess the impact of silencing LINC00961 transcript on endothelial function, we depleted LINC00961 levels in HUVECs by 70%, utilising dsRNAs (Figure 3A). In an in vitro 2D Matrigel tubule network formation assay, LINC00961 silencing resulted in attenuated branch formation (Figure 3B & C). Calcein AM was used to confirm that the lack of branch formation following LINC00961 depletion was not a consequence of apoptosis (Figure 3C). We confirmed that LINC00961 silencing did not affect cell viability using Annexin V and PI staining (Supplementary Figure 3A). We then replicated the network formation phenotype via a GapmeR depletion strategy (Supplementary Figure 4). Moreover, silencing LINC00961 led to a significant reduction in cell adhesion (Figure 3D) and endothelial membrane barrier integrity (Figure 3E & Supplementary Figure 3B). We also observed a trend towards a reduction in cell proliferation (Supplementary Figure 3C), and migration (Supplementary Figure 3D). To investigate whether LINC00961 played a cis-regulatory role in the expression of the closely located gene HRCT1, we tested HRCT1 transcript levels in siRNA LINC00961 depleted cells. qRT-PCR analysis showed that HRCT1 expression was unaltered by LINC00961 modulation (Supplementary Figure 5A & B). Similarly, siRNA silencing of HRCT1 did not affect LINC00961 levels (Supplementary Figure 5C & D).

**Murine LINC00961/SPAAR locus knock out reduces adductor muscle capillary density following hind limb ischaemia (HLI)**

To assess the role of the LINC00961 locus in vivo, we established a knock out (KO) mouse where the entire locus was deleted (Figure 4A). We first confirmed the absence of the LINC00961 mouse transcript by qRT-PCR (Supplementary Figure 6A). We then tested the efficacy of injury-induced angiogenesis compared to littermate controls at two time points. After 7 days, the capillary density between KO and WT animals was not significantly altered in the non-ischaemic leg (p=0.2471). However, at 7 days after HLI LINC00961−/− mice had a lower capillary density in the ischaemic adductor muscle compared to controls (Figure 4B). This was therefore comparable to the in vitro tubule
formation data in LINC00961 depleted HUVECs. Interestingly, KO animals had a significant decrease in the number of α-smooth muscle actin (αSMA) positive vessels at baseline compared to WT animals but this difference was not evident after injury (Figure 4C). We also analysed Laser Doppler ratio, capillary density and αSMA positive vessels at 21 days. No significant differences at this later time point were observed (Supplementary Figure 6).

The LINC00961 locus encodes a biologically functional RNA
We next investigated the angiogenic effect of overexpressing either the full length LINC00961 transcript or the SPAAR ORF sequence in HUVECs, using lentiviral vectors (LV) (Figure 5A). We also generated a LV-ΔΔATG961 construct (Figure 5A), corresponding to the full length transcript with mutations in the ORF initiation codons to block translation. qRT-PCR (Figure 5B) and western blotting (Figure 5C) confirmed overexpression. Overexpression of the LV-SPAAR construct significantly enhanced endothelial network formation, whereas LV-ΔΔATG961 produced opposite results, significantly inhibiting angiogenesis (Figure 5D & E). These data showed that the production of SPAAR induces network formation, whereas the LINC00961 RNA alone possesses an inhibitory effect, independent of SPAAR micropeptide production, thus unveiling a bona fide IncRNA function for the LINC00961 RNA. Furthermore, we observed that LV-mediated overexpression of SPAAR, but not the LINC00961 transcript, reduced endothelial barrier integrity (Figure 5F). As cellular localisation of IncRNA transcripts is informative with regards to their associated mechanisms, we determined the subcellular localisation of LINC00961 using RNA-fluorescent in situ hybridisation (FISH) (Supplementary Figure 8A & B) and cell fractionation (Supplementary Figure 8C) and showed the presence of LINC00961 in both the nucleus and the cytoplasm.

Identification of binding partners for LINC00961 RNA and SPAAR micropeptide
As both LINC00961 and SPAAR are functionally relevant for ECs, we used RNA and protein pull-downs combined with mass spectrometry to identify the protein binding partners of the IncRNA and SPAAR micropeptide in HUVECs (Figure 6). 417 proteins were found in the LINC00961 pulldown samples, which were not in the pulldown with the beads alone or the control GFP RNA (Figure 6B & Supplementary Tables 3 and 4). GO analysis showed enrichment of terms related to cell-cell adhesion and cortical actin arrangement (Figure 6D). The top candidate was the G-actin sequestering molecule, thymosin beta 4-x (Tβ4) which is associated with reorganisation of the actin cytoskeleton 30 and is also involved in angiogenesis 30, 31. Tβ4 functions within an actin organisation pathway with other actin associated molecules including Cofilin-1 and Profilin-1 32. Both Profilin-1 and Cofilin-1 were enriched in the LINC00961 immunoprecipitation (Supplementary Table 3); suggesting LINC00961 may play a role in actin cytoskeleton remodelling. To confirm the interaction between LINC00961 and Tβ4, we carried out immunoprecipitation of endogenous Tβ4 protein in HUVECs. qRT-PCR confirmed the detection of LINC00961 in Tβ4 immunoprecipitation samples, thus independently validating an interaction of LINC00961 with Tβ4 (Supplementary Figure 10A). Immunofluorescence of Tβ4 in HUVECs confirmed the presence of Tβ4 in the cytoplasm in accordance with a plausible interaction with LINC00961 (Supplementary Figure 10B).

We next identified protein binding partners for SPAAR. We found 40 proteins enriched in the HA-SPAAR pull-down compared to the IgG pull-down controls and compared to the pull-downs in control cells not expressing the fusion protein (Supplementary Table 5 & 6). GO analysis of SPAAR targets showed enrichment of terms related to immunity (Figure 6E). SPAAR has been previously shown to bind the v-ATPase complex in HEK293 33. However, these proteins were not found in the SPAAR pull-down in HUVECs, suggesting a different function for SPAAR in ECs. The top hit for SPAAR interactors was SYNE1, also known as NESPRIN-1, a regulator of EC shape and migration 34.

Thymosin beta 4-x depletion phenocopies LV-ΔΔATG961 overexpression
To characterise the function of LINC00961 and Tβ4 interaction, we assessed whether they co-regulated each other’s expression. siRNA silencing of TMSB4X (Supplementary Figure 9A) did not alter LINC00961 transcript levels (Supplementary Figure 9B). Similarly, silencing LINC00961 or overexpressing LV-ΔΔATG961 did not change TMSB4X transcript levels (Supplementary Figure 9C & 9D). The known pro-angiogenic effect of Tβ4 30, 31 was confirmed in our system, with a 49±16%
reduction in network formation following TMSB4X depletion (Figure 7A & 7B). This reduction is similar to the overexpression of LINC00961 transcript without the production of SPAAR micropeptide (LV-ΔΔATG961), suggesting that LINC00961 IncRNA might negatively regulate Tβ4-mediated angiogenesis.

(vi) DISCUSSION

Using RNA-seq, we identified LINC00961 as an endothelial enriched transcript. The strong impact on the endothelial phenotype following LINC00961 level manipulations confirmed the relevance of our candidate selection using the combination of our hESC to EC RNA-seq with ENCODE RNA-seq datasets. This further highlights the need to investigate the role of IncRNA transcripts in endothelial biology.

In this study, we provide in vitro and in vivo evidence that the LINC00961 locus has a function in ECs. Whilst siRNA KD in vitro affects many aspects of EC biology (angiogenesis, adhesion, proliferation, migration, and membrane integrity), we assessed the angiogenic role in a murine KO model. LINC00961−/− mice had fewer αSMA positive vessels at day 7 baseline, suggesting a defect in the development, maturity and or stability of larger vessels. After injury, KO mice has fewer capillaries at day 7, indicating a reduced capacity of the endothelium to undergo angiogenesis after injury. However, the effect of the KO was not observed by day 21 post HLI. This suggests the KO animals may have a slower recovery rate after injury (due to an impairment in EC function), or activate compensatory mechanisms to maintain vessel numbers after injury. As we have a global KO, we cannot exclude the contribution of LINC00961 deletion in other cell types to this phenotype. To further investigate the role of this locus in EC behaviour, it would be worthwhile to switch to an EC-specific and conditional LINC00961 KO mouse model. In addition, it would be interesting to assess the effect of LINC00961 deletion in early development of vessel establishment and further characterise the dynamics of vessel recovery early in the HLI model.

Previous studies have outlined the role of the micropeptide SPAAR, encoded from the LINC00961 locus, during muscle regeneration 23. In our study, we showed opposing roles of LINC00961 RNA and SPAAR micropeptide in angiogenesis, one being anti- and the other pro-angiogenic respectively. The reduction in endothelial barrier integrity with SPAAR overexpression further validates our hypothesis that SPAAR is pro-angiogenic. Indeed, plastic junctions are required for sprouting angiogenesis 35. It would therefore be interesting to test the permeability of new SPAAR induced vessels in an animal system using a plasma tracer.

To our knowledge, this is the first reported bi-functional locus in a cardiovascular setting. In other biological contexts, loci producing protein or functional ncRNAs through alternative splicing have been described 36–38. The novelty of the LINC00961 locus is that the SPAAR micropeptide is produced from the functional LINC00961 RNA instead of an alternative splicing transcript without an ORF. This configuration implies the requirement of a regulatory mechanism to control the levels of LINC00961 RNA and SPAAR micropeptide independently of each other. The switch between LINC00961 and SPAAR could be controlled at the translation level, similarly to the STORM micropeptide whose translation initiation is regulated by eIF4E phosphorylation 39. However, the functional activity of the IncRNA encoding the STORM micropeptide has never been demonstrated. Expression of the LINC00961 transcript is high in basal HUVECs and detectable by qRT-PCR, in contrast, we are only
able to see the presence of SPAAR micropeptide in LV-SPAAR conditions. This limitation is likely due to either very low protein levels in basal HUVECs or the detection limit of the antibody. The precise molecular control of LINCO0961 transcript and SPAAR levels needs further dissection in light of these findings.

We show that LINCO0961 RNA binds Tβ4, a well-established actin binding protein with many additional functions including anti-inflammatory and anti-apoptotic properties, and a role in cell migration and angiogenesis. As TMSB4X transcript levels were not affected by LINCO0961 depletion, we propose that LINCO0961 regulates Tβ4 protein function. The enrichment of Profilin-1 and Cofilin-1, actin monomer binding proteins, in the LINCO0961 immuno-precipitation suggests a potential role for LINCO0961 in actin recycling. Like Tβ4, Profilin-1 sequesters G-actin maintaining a large pool of monomeric actin. Unlike Tβ4 however, the high affinity of Profilin-1 for ATP allows it to act as a catalyst for the conversion of G-actin. ADP to G-actin. ATP, hence aiding the polymerisation of G-actin to F-actin filaments. In fact, Profilin-1, Tβ4, and actin have been shown to produce a complex.

This, alongside the fact that Cofilin-1 and Tβ4 have been shown to co-localise in multiple cells types, further validates the nature of their finely balanced roles in cytoskeletal dynamics. It would be of interest to dissect the interactions of these three proteins with LINCO0961 in future.

We show that SPAAR binds to SYNE1, another actin binding protein, which suggests that the pro-angiogenic effects of SPAAR could be mediated through SYNE1 and the actin cytoskeleton. This is in contrast to our proposed mechanism of action of LINCO0961, which may negatively affect actin cytoskeleton rearrangement through interaction with Tβ4. SYNE1 is involved in the cellular organisation of organelles via connecting them to the actin cytoskeleton. SYNE1 is especially important as a member of the linker of nucleoskeleton and cytoskeleton complex which tethers the nuclear lamina to the actin cytoskeleton during nuclear positioning and cell polarisation. Interestingly, SYNE1 is highly expressed in skeletal and cardiac muscle cells as it is essential in maintaining the characteristic peripherally located nuclei. Matsumoto and colleagues (2017) describe rapid muscle regeneration in mice lacking SPAAR; it would be interesting to ascertain if this phenomenon is in part mediated by an interaction, or lack thereof, between SPAAR and SYNE1. Furthermore, SYNE1 siRNA KD in HUVECs has been shown to reduce tube formation in a Matrigel assay and decreased migration, similar to our results with KD of the LINCO0961/SPAAR locus. Cytoskeletal remodelling is a dynamic process which is constantly being influenced by internal and external signals, with many actin binding proteins having been identified. Here, we show that LINCO0961 and SPAAR have independent actin binding protein partners that could influence downstream cytoskeletal architecture. It will be of interest to investigate if, and how, IncRNA and micropeptide levels can change cellular behaviour through cytoskeletal changes.

In conclusion, our study provides important evidence for the expression and function of LINCO0961 in ECs. Our work shows a role for the LINCO0961 RNA, independent of the micropeptide SPAAR. This highlights the importance of a detailed bioinformatic and experimental approach to reveal the contribution of putative IncRNAs and their encoded proteins in cell behaviours.

(vii) FUNDING

The British Heart Foundation supported this work (Program grants: RG/14/3/30706 to AHB, RG/15/5/31446 to CE and RG/17/4/32662 to AMR and project grant and FS/17/27/32698 to AHB). Professor Baker is supported by EU CARDIOREGENIX, The British Heart Foundation Chair of Translational Cardiovascular Sciences (CH/11/2/28733), European Research Council (EC 338991 VASCIMIR). AHB, MB, AMR and CE are all supported by the BHF Regenerative Medicine Centre (RM/13/2/30158). MB is supported by the British Heart Foundation (FS/16/4/31831). MB is further supported by the BHF Centre for Vascular Regeneration (RM/17/3/3381).
ACKNOWLEDGMENTS

Flow cytometry data was generated with support from the QMRI Flow Cytometry and cell sorting facility, University of Edinburgh. Mass spectrometry data was generated with support from the IGMM Mass Spectrometry facility, University of Edinburgh. Animal experiments were supported by the BVS facility, University of Edinburgh.

We thank G. Aitchison, Y. Harcus, K. Newton, O. Kelepouri, and L. Rose for technical assistance.

CONFLICT OF INTEREST

None declared.

REFERENCES


(xi) FIGURE LEGENDS

**Figure 1:** Identification of endothelial cell enriched genes. (A) Schematic representation of the RNA-seq samples: day 0 H9 hESC (ESC); Day 3 mesodermal population CD326^low^CD56^+^ (MP); Day 3 remaining population (non-MP); Day 7 EC CD144^+^CD31^+^ (EC); Day 7 remaining population (non-EC); Human Saphenous vein endothelial cell (HSVEC). (B) Principle component analysis (PCA) of the RNA-seq samples. The plot was generated on the regularized log transformed data using DESeq2. (C) Summary of the selection of candidates to identify genes enriched in ‘immature’ and ‘mature’ ECs (D) Heatmap showing the expression data (as row z-score of the Log2(FPKM+1)) during differentiation of the 278 EC enriched genes. (E) Heatmap showing the expression data (as row z-score of the Log2(FPKM+1)) of the 278 EC enriched genes in ENCODE RNA-seq samples.

**Figure 2:** LINC00961 is enriched in immature and mature endothelial cells. (A) Heatmap of the 6 lncRNAs identified in our EC differentiation protocol in each of the isolated cell populations. (B) Heatmap of these 6 lncRNAs in ENCODE RNA-seq samples including various types of EC lineages such as, venous, arterial and lymphatic ECs. (C) Genomic organisation of the LINC00961 gene, read profile from the ESC to EC RNA-seq and conservation track based on UCSC alignment and PhyloP score.
Figure 3: Functional impact of LINC00961/SPAAR depletion in endothelial cells. (A) Confirmation of the dsiRNA-mediated depletion of LINC00961 transcript in HUVECs by qRT-PCR (n=4, unpaired t-test). (B) Network formation assay in LINC00961 depleted HUVECs. Branch length assessed by Image J Angiogenesis plugin. (n=3, unpaired t-test). (C) Representative phase contrast and Calcein AM staining of network formation assay of LINC00961 depleted and control HUVECs. Phase Scale bar =0.5mm. Calcein AM Scale bar =0.1mm. (D) Impact of LINC00961 depletion on HUVEC adhesion (n=3). (E) Analysis of average barrier resistance, expressed as Rb [Ohm x cm²], across a 10 hour time course (n=4 except for mock n=3, one-way ANOVA). For data represented as fold change, the statistical analysis was done on the Log2 Fold Change using a One Sample t-test. On the graphs, *p<0.05 **p<0.01 ***p<0.001.

Figure 4: LINC00961/SPAAR KO mice have a reduced adductor muscle capillary density following hind limb ischaemia at 7 days. (A) Schematic representation of the deleted region of the LINC00961 mouse locus using CRISPR/Cas9 technology by Taconic®. Red arrows indicate the position of the guide RNA strands utilised to delete the whole locus. (B) Capillary density per sample. Five random regions of interest from 3 sections per sample were counted (n= 4 wild type mice/6 knock out mice, one-way ANOVA, ** p<0.01, ns - not significant). (C) α-smooth muscle actin (αSMA) positive vessel density per sample. (D) Representative adductor muscle immunofluorescent images: Isolectin b4 (IB4) capillary/endothelium, αSMA, and nuclear DAPI, scale bar 50μm. Zoomed panel on left corresponds to red box on area of WT control limb image.

Figure 5: Impact of LINC00961 transcript and SPAAR micropeptide overexpression in in vitro angiogenic assays. (A) Schematic representation of LINC00961 LV constructs with transcript length in base pairs (bp) and encoded peptide length in amino acids (aa). (B) qRT-PCR validation of the LV constructs overexpression in HUVECs using primers targeting the ORF sequence. Unpaired t-test, comparison test versus LV-EMPTY (n=4). (C) Representative western blot of SPAAR micropeptide and β-actin in HUVECs infected with the LV constructs. (D) Network formation assay comparing HUVECs transfected with LV constructs. Branch length assessed by Image J Angiogenesis plugin. Unpaired t-test versus LV-EMPTY, (n=3). (E) Representative Phase contrast of network formation assay of HUVECs transfected with LV constructs. (F) Analysis of average barrier resistance, expressed as Rb [Ohm x cm²], across a 10 h time course (n=4, one-way ANOVA). Scale bar =0.5mm. On the graphs, *p<0.05, **p<0.01, ***p<0.001.

Figure 6: LINC00961 and SPAAR both bind to actin binding proteins. (A) Schematic of the LINC00961 RNA and SPAAR peptide pull-down experiments in HUVECs. (B) List of the top 10 proteins identified in LINC00961 RNA pull-down (ranked on label free quantification value) (C) List of the top 10 proteins identified in HA-SPAAR peptide pull-down (ranked on label free quantification value) (D) GO analysis on enriched proteins from LINC00961 immunoprecipitation. (E) GO analysis on enriched proteins from SPAAR immunoprecipitation.

Figure 7: Thymosin beta 4-x KD in HUVECS has a similar phenotype to LV-ΔΔATG961 overexpression on tubule formation. (A) Network formation assay in dsiRNA-mediated TMSB4X depleted HUVECs. Branch length assessed by Image J Angiogenesis plugin, n=5, unpaired t-test. (B) Representative phase contrast and Calcein AM staining of network formation assay of depleted HUVECs. Phase contrast Scale bar =0.5mm. Calcein AM Scale bar = 0.1 mm. On the graphs, *p<0.05, **p<0.01, ***p<0.001.
A. hESC d0
   Stemline+Cytokine
   EGM-2+VEGF
   MP
   CD326+CD56+
   d3
   EBs
   CD326low
   CD56+
   non-MP
   CD326+CD56-
   EC
   CD144+
   CD31+
   d7
   non-EC
   CD144-
   HSVEC

B. PC2: 23% variance
   PC1: 57% variance
   ESC
   MP
   non-MP
   non-EC
   HSVEC

C. Identification of genes enriched in 'immature' and 'mature' Endothelial cells
   ESC to EC RNAseq
   | Enriched in day7 EC | Up-regulated in HSVEC |
   | 778 genes          |                      |
   = ENCODE RNAseq
   | Expressed in ENCODE |
   | EC 409 genes       |
   | Enriched in ENCODE |
   | EC versus SMC 278 genes |

D. Heatmap
   ESC
   MP
   non MP
   EC
   non EC
   HSVEC

E. Heatmap
   Heatmap of gene expression data.
Figure(s)
Figure(s)
Identification of genes enriched in ‘immature’ and ‘mature’ Endothelial cells

**ESC to EC RNAseq**
- Enriched in day7 EC: 778 genes
- Up-regulated in HSVEC: 409 genes

**ENCODE RNAseq**
- Expressed in ENCODE EC: 382 genes
- Enriched in ENCODE EC versus SMC: 278 genes
A

Murine LINC00961/SPAAR (5430416O09Rik)
Chromosome 4 predicted 4.086Kb

B

Capillary Density (capillaries/mm²)

C

αSMA +ve vessels (per mm²)

D

Wild type
Control limb

Wild type
Ischaemic limb

LINC00961−/−
Control limb

LINC00961−/−
Ischaemic limb

αSMA
Isolectin b4
DAPI

50µm
### A

- **LINC00961**
  - 1612bp
  - 75aa ORF

- **ΔΔATG961**
  - 1612bp
  - No ORF

- **SPAAR**
  - 228bp
  - 75aa ORF

### B

**LV-Construct expression**

- Relative LV-Construct expression normalised to UBC

### C

**Network formation assay**

- UT
- LV-EMPTY
- LV-LINC00961
- LV-ΔΔATG961
- LV-SPAAR

- SPAAR 10kDa
- ACTB 42kDa

### D

**Total branch length (mm)**

- UT
- LV-EMPTY
- LV-LINC00961
- LV-ΔΔATG961
- LV-SPAAR

### E

- LV-EMPTY
- LV-LINC00961
- LV-SPAAR
- LV-ΔΔATG961

### F

**Rb (Ohm X cm^2)**

- LV-EMPTY
- LV-ΔΔATG961
- LV-SPAAR

**EC barrier integrity**

- ***
- **
- *
- ns
A

LINC00961 RNA pulldown

Biotinylated LINC00961 → HUVEC lysate → Proteins → Streptavidin bead → Mass Spectrometry analysis

SPAAR peptide pulldown

Lysate of HUVEC overexpressing HA-SPAAR → Anti-HA bead → Mass Spectrometry analysis

B

MARCKS, CALR, RPF2, MSI2, PRKCDDBP, APEX1, FAM50A, EIF4H, DAZAP1, TMSB4X

Label Free Quantification Intensity

C

C1QB, IGKV2D-29, ZCCHC17, MATN2, GPI, DCD, SQSTM1, DNAH12, SPAAR, SYNE1

Label Free Quantification Intensity

D

GO Terms: LINC00961 RNA pulldown

Cell-cell adhesion, Formation of translation preinitiation complex, Cortical actin cytoskeleton organization, Hippo signaling, Regulation of translational initiation

E

GO Terms: SPAAR peptide pulldown

Humoral immune response, Leukocyte mediated immunity, Proteolysis, Response to axon injury
Network formation assay

A

B

Mock
Si Control
Si TMSB4X #1
Si TMSB4X #2

Total branch length (mm)

*
**Graphical Abstract**

LINC00961 is located at the **LINC00961 locus**. It undergoes translation to produce **SPAAR** and **TMSB4X**. SPAAR exhibits a **non-coding function** and a **micropeptide function**, while TMSB4X is located in the **cytoplasm** and **nucleus**. The balance between SPAAR and TMSB4X is crucial in determining the angiogenic vs anti-angiogenic properties of the **endothelial cell**. The diagram highlights the role of the **SYNE1** gene in this process.