Edinburgh Research Explorer

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

(i)   <u>Cover Letter</u>

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.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.

(ii) **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-sma: 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

2

lincRNA: Intergenic lncRNA
lncRNA: 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
Tβ4: 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 [1]. Aberrant activation however, leads to EC dysfunction that can cause systemic vascular pathology [1, 2]. 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 [3, 4].

Several groups have demonstrated the ability to differentiate ECs from human embryonic stem cells (hESC) [5-7]. This protocol yields ECs that are relatively immature and express genes that are somewhat distinct from those of mature ECs from various vascular beds [8], 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 [7] and *in vivo,* by their ability to improve vascular density and perfusion in a murine model of hind limb ischaemia (HLI) [9]. 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 [10]. 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, lincRNAs are intergenic lncRNAs with no overlap with protein-coding genes. While some lincRNAs 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 [11]. LncRNAs show spatio-temporal expression, and are poorly conserved between species [12] , 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 [13, 14]. In the vascular endothelium, TIE1-AS1 was the first described endothelial specific lncRNA, involved in modulating TIE-1 expression and regulating endothelial vessel formation [15]. A comprehensive transcriptome analysis of early cardiovascular development revealed the regulation of several lncRNAs and led to the characterisation of ALIEN and PUNISHER [16]. Recently, the hypoxia-induced lncRNA, GATA6-AS, was shown to epigenetically regulate angiogenesis through its interaction with the epigenetic regulator LOXL2 [17, 18].

3

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 [19, 20]. For example, LINC00948 has been reclassified as a protein-coding gene, as it encodes myoregulin, which inhibits the calcium ATPase SERCA in muscle [21]. Similarly, the micropeptide DWORF encoded by lncRNA NONMMUG026737 activates the SERCA pump [22]. Noteworthy to this study, a conserved micropeptide termed small regulatory polypeptide of amino acid response (SPAAR) was recently shown to be encoded by the *LINC00961* locus [23]. SPAAR attenuates lysosomal v-ATPases interaction with mTORC1 under amino acid stimulation and modulates skeletal muscle regeneration following cardiotoxin injury [23]. 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[24], 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-x (Tβ4). 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 (HUVEC) 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 [25]. RNA-Seq analysis was performed as previously described [26] with minor modifications. Ensembl GRCh38 was used for transcriptome annotation. Read counts for each gene were obtained using HTSeq [27]. The differential expression was analysed using DESeq2 [28]. 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 FITC 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. LINC00961[-/-] 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 LINC00961[-/-] 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 lncRNA, obtained using the T7 RiboMAX Express Large Scale RNA Production System (Promega, UK). The biotinylated lncRNA 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 dCt analysis while the statistical analyses were done on dCt values. For data represented as fold change, the statistical analysis was done on the Log2 Fold Change using a One Sample t-test.

(v)   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.

5

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 [24], 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 [23]. 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 dsiRNAs (**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* lncRNA 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 lncRNA 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 lncRNA and SPAAR micropeptide in HUVECs **(Figure 6)**. 147 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 [29] 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* lncRNA 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 lncRNA 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 LINC00961deletion 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 lncRNA 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 *LINC00961* transcript and SPAAR levels needs further dissection in light of these findings.

We show that *LINC00961* 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 [29]. As *TMSB4X* transcript levels were not affected by *LINC00961* depletion, we propose that *LINC00961* regulates Tβ4 protein function. The enrichment of Profilin-1 and Cofilin-1, actin monomer binding proteins, in the *LINC00961* immuno-precipitation suggests a potential role for *LINC00961* 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 [40]. In fact, Profilin-1, Tβ4, and actin have been shown to produce a complex [41]. 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 [42],[43]. It would be of interest to dissect the interactions of these three proteins with *LINC00961* 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 *LINC00961*, 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 [44]. Interestingly, SYNE1 is highly expressed in skeletal and cardiac muscle cells as it is essential in maintaining the characteristic peripherally located nuclei [45]. 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 tubule formation in a Matrigel assay and decreased migration [34, 46], similar to our results with KD of the LINC00961/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 [47]. Here, we show that *LINC00961 and* SPAAR have independent actin binding protein partners that could influence downstream cytoskeletal architecture. It will be of interest to investigate if, and how, lncRNA and micropeptide levels can change cellular behaviour through cytoskeletal changes.

In conclusion, our study provides important evidence for the expression and function of *LINC00961* in ECs. Our work shows a role for the *LINC00961* RNA, independent of the micropeptide SPAAR. This highlights the importance of a detailed bioinformatic and experimental approach to reveal the contribution of putative lncRNAs and their encoded proteins in cell behaviours.

9

(viii)   ACKNOWLEDGMENTS

(ix)   CONFLICT OF INTEREST
None declared.

(x)   REFERENCES

1. Toborek M, Kaiser S. Endothelial cell functions.¶ Relationship to atherogenesis. *Basic research in cardiology* 1999;**94**:295-314.
2. De Vries MR, Simons KH, Jukema JW, Braun J, Quax PH. Vein graft failure: from pathophysiology to clinical outcomes. *Nature Reviews Cardiology* 2016;**13**:451.
3. Pearson J. Normal endothelial cell function. *Lupus* 2000;**9**:183-188.
4. Su JB. Vascular endothelial dysfunction and pharmacological treatment. *World journal of cardiology* 2015;**7**:719.
5. Zhao H, Zhao Y, Li Z, Ouyang Q, Sun Y, Zhou D, Xie P, Zeng S, Dong L, Wen H. FLI1 and PKC co-activation promote highly efficient differentiation of human embryonic stem cells into endothelial-like cells. *Cell death & disease* 2018;**9**:131.
6. Rufaihah AJ, Huang NF, Jamé S, Lee JC, Nguyen HN, Byers B, De A, Okogbaa J, Rollins M, Reijo-Pera R. Endothelial cells derived from human iPSCS increase capillary density and improve perfusion in a mouse model of peripheral arterial disease. *Arteriosclerosis, thrombosis, and vascular biology* 2011;**31**:e72-e79.
7. Boulberdaa M, Scott E, Ballantyne M, Garcia R, Descamps B, Angelini GD, Brittan M, Hunter A, McBride M, McClure J. A role for the long noncoding RNA SENCR in commitment and function of endothelial cells. *Molecular Therapy* 2016;**24**:978-990.
8. McCracken IR, Taylor RS, Kok FO, de la Cuesta F, Dobie R, Henderson BE, Mountford JC, Caudrillier A, Henderson NC, Ponting CP. Transcriptional dynamics of pluripotent stem cell-derived endothelial cell differentiation revealed by single-cell RNA sequencing. *European heart journal* 2019.
9. MacAskill MG, Saif J, Condie A, Jansen MA, MacGillivray TJ, Tavares AA, Fleisinger L, Spencer HL, Besnier M, Martin E. Robust Revascularization in Models of Limb Ischemia Using a Clinically Translatable Human Stem Cell-Derived Endothelial Cell Product. *Molecular Therapy* 2018.
10. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng ZP, Snyder M, Dermitzakis ET, Stamatoyannopoulos JA, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SCJ, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Dutta A, Guigo R, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng DY, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Flicek P, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermuller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT,

Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung WK, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu JQ, Lian Z, Lian J, Newburger P, Zhang XQ, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Dermitzakis ET, Margulies EH, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei CL, Ruan YJ, Snyder M, Birney E, Struhl K, Gerstein M, Antonarakis SE, Gingeras TR, Brown JB, Flicek P, Fu YT, Keefe D, Birney E, Denoeud F, Gerstein M, Green ED, Kapranov P, Karaoz U, Myers RM, Noble WS, Reymond A, Rozowsky J, Struhl K, Siepel A, Stamatoyannopoulos JA, Taylor CM, Taylor J, Thurman RE, Tullius TD, Washietl S, Zheng DY, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Collins FS, Margulies EH, Cooper GM, Asimenos G, Thomas DJ, Dewey CN, Siepel A, Birney E, Keefe D, Hou MM, Taylor J, Nikolaev S, Montoya-Burgos JI, Loytynoja A, Whelan S, Pardi F, Massingham T, Brown JB, Huang HY, Zhang NR, Bickel P, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA, Gerstein M, Antonarakis SE, Batzoglou S, Goldman N, Hardison RC, Haussler D, Miller W, Pachter L, Green ED, Sidow A, Weng ZP, Trinklein ND, Fu YT, Zhang ZDD, Karaoz U, Barrera L, Stuart R, Zheng DY, Ghosh S, Flicek P, King DC, Taylor J, Ameur A, Enroth S, Bieda MC, Koch CM, Hirsch HA, Wei CL, Cheng J, Kim J, Bhinge AA, Giresi PG, Jiang N, Liu J, Yao F, Sung WK, Chiu KP, Vega VB, Lee CWH, Ng P, Shahab A, Sekinger EA, Yang A, Moqtaderi Z, Zhu Z, Xu XQ, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Clelland GK, Wilcox S, Dillon SC, Andrews RM, Fowler JC, Couttet P, James KD, Lefebvre GC, Bruce AW, Dovey OM, Ellis PD, Dhami P, Langford CF, Carter NP, Vetrie D, Kapranov P, Nix DA, Bell I, Patel S, Rozowsky J, Euskirchen G, Hartman S, Lian J, Wu JQ, Urban AE, Kraus P, Van Calcar S, Heintzman N, Kim TH, Wang K, Qu CX, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Zhang XL, Xu MS, Haidar JNS, Yu Y, Birney E, Weissman S, Ruan YJ, Lieb JD, Iyer VR, Green RD, Gingeras TR, Wadelius C, Dunham I, Struhl K, Hardison RC, Gerstein M, Farnham PJ, Myers RM, Ren B, Snyder M, Thomas DJ, Rosenbloom K, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakkapallayil A, Barber G, Kuhn RM, Karolchik D, Haussler D, Kent WJ, Dermitzakis ET, Armengol L, Bird CP, Clark TG, Cooper GM, de Bakker PIW, Kern AD, Lopez-Bigas N, Martin JD, Stranger BE, Thomas DJ, Woodroffe A, Batzoglou S, Davydov E, Dimas A, Eyras E, Hallgrimsdottir IB, Hardison RC, Huppert J, Sidow A, Taylor J, Trumbower H, Zody MC, Guigo R, Mullikin JC, Abecasis GR, Estivill X, Birney E, Bouffard GG, Guan XB, Hansen NF, Idol JR, Maduro VVB, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang HY, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu BL, de Jong PJ, Consortium EP. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007;**447**:799-816.

11. Ransohoff JD, Wei Y, Khavari PA. The functions and unique features of long intergenic non-coding RNA. *Nature reviews Molecular cell biology* 2018;**19**:143.

12. Ulitsky I. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nat Rev Genet* 2016;**17**:601-614.

13. Deng L, Bradshaw AC, Baker AH. Role of noncoding RNA in vascular remodelling. *Curr Opin Lipidol* 2016;**27**:439-448.

14. Uchida S, Dimmeler S. Long noncoding RNAs in cardiovascular diseases. *Circ Res* 2015;**116**:737-750.

15. Maisonpierre PC, Suri C, Jones PF, Bartunkova S, Wiegand SJ, Radziejewski C, Compton D, McClain J, Aldrich TH, Papadopoulos N, Daly TJ, Davis S, Sato TN, Yancopoulos GD. Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis. *Science* 1997;**277**:55-60.

16. Kurian L, Aguirre A, Sancho-Martinez I, Benner C, Hishida T, Nguyen TB, Reddy P, Nivet E, Krause MN, Nelles DA, Esteban CR, Campistol JM, Yeo GW, Belmonte JCI. Identification of

Novel Long Noncoding RNAs Underlying Vertebrate Cardiovascular Development. *Circulation* 2015;**131**:1278-1290.

17. Neumann P, Jae N, Knau A, Glaser SF, Fouani Y, Rossbach O, Kruger M, John D, Bindereif A, Grote P, Boon RA, Dimmeler S. The lncRNA GATA6-AS epigenetically regulates endothelial gene expression via interaction with LOXL2. *Nat Commun* 2018;**9**:237.

18. Monteiro JP, Bennett M, Rodor J, Caudrillier A, Ulitsky I, Baker AH. Endothelial Function and Dysfunction in the Cardiovascular System: The Long Non-Coding Road. *Cardiovascular Research* 2019.

19. Andrews SJ, Rothnagel JA. Emerging evidence for functional peptides encoded by short open reading frames. *Nat Rev Genet* 2014;**15**:193-204.

20. Makarewich CA, Olson EN. Mining for Micropeptides. *Trends in Cell Biology* 2017;**27**:685-696.

21. Anderson DM, Anderson KM, Chang CL, Makarewich CA, Nelson BR, McAnally JR, Kasaragod P, Shelton JM, Liou J, Bassel-Duby R, Olson EN. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* 2015;**160**:595-606.

22. Nelson BR, Makarewich CA, Anderson DM, Winders BR, Troupes CD, Wu F, Reese AL, McAnally JR, Chen X, Kavalali ET, Cannon SC, Houser SR, Bassel-Duby R, Olson EN. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* 2016;**351**:271-275.

23. Matsumoto A, Pasut A, Matsumoto M, Yamashita R, Fung J, Monteleone E, Saghatelian A, Nakayama KI, Clohessy JG, Pandolfi PP. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature* 2017;**541**:228-+.

24. van Heesch S, Witte F, Schneider-Lunitz V, Schulz JF, Adami E, Faber AB, Kirchner M, Maatz H, Blachut S, Sandmann C-L. The translational landscape of the human heart. *Cell* 2019.

25. Boulberdaa M, Scott E, Ballantyne M, Garcia R, Descamps B, Angelini GD, Brittan M, Hunter A, McBride M, McClure J, Miano JM, Emanueli C, Mills NL, Mountford JC, Baker AH. A Role for the Long Noncoding RNA SENCR in Commitment and Function of Endothelial Cells. *Mol Ther* 2016;**24**:978-990.

26. Ballantyne MD, Pinel K, Dakin R, Vesey AT, Diver L, Mackenzie R, Garcia R, Welsh P, Sattar N, Hamilton G, Joshi N, Dweck MR, Miano JM, McBride MW, Newby DE, McDonald RA, Baker AH. Smooth Muscle Enriched Long Noncoding RNA (SMILR) Regulates Cell Proliferation. *Circulation* 2016;**133**:2050-2065.

27. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;**31**:166-169.

28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;**15**:550.

29. Kuzan A. Thymosin beta as an Actin-binding Protein with a Variety of Functions. *Adv Clin Exp Med* 2016;**25**:1331-1336.

30. Smart N, Risebro CA, Melville AA, Moses K, Schwartz RJ, Chien KR, Riley PR. Thymosin beta4 induces adult epicardial progenitor mobilization and neovascularization. *Nature* 2007;**445**:177-182.

31. Smart N, Rossdeutsch A, Riley PR. Thymosin beta4 and angiogenesis: modes of action and therapeutic potential. *Angiogenesis* 2007;**10**:229-241.

32. Skruber K, Read TA, Vitriol EA. Reconsidering an active role for G-actin in cytoskeletal regulation. *J Cell Sci* 2018;**131**.

33. Matsumoto A, Pasut A, Matsumoto M, Yamashita R, Fung J, Monteleone E, Saghatelian A, Nakayama KI, Clohessy JG, Pandolfi PP. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature* 2016.

34. Chancellor T, Lee J, Thodeti CK, Lele T. Actomyosin tension exerted on the nucleus through nesprin-1 connections influences endothelial cell adhesion, migration, and cyclic strain-induced reorientation. *Biophysical journal* 2010;**99**:115-123.

35. Gordon EJ, Fukuhara D, Weström S, Padhan N, Sjöström EO, van Meeteren L, He L, Orsenigo F, Dejana E, Bentley K. The endothelial adaptor molecule TSAd is required for VEGF-induced angiogenic sprouting through junctional c-Src activation. *Sci Signal* 2016;**9**:ra72-ra72.

36. Caretti G, Schiltz RL, Dilworth FJ, Di Padova M, Zhao P, Ogryzko V, Fuller-Pace FV, Hoffman EP, Tapscott SJ, Sartorelli V. The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation. *Dev Cell* 2006;**11**:547-560.

37. Hube F, Velasco G, Rollin J, Furling D, Francastel C. Steroid receptor RNA activator protein binds to and counteracts SRA RNA-mediated activation of MyoD and muscle differentiation. *Nucleic Acids Res* 2011;**39**:513-525.

38. Grelet S, Link LA, Howley B, Obellianne C, Palanisamy V, Gangaraju VK, Diehl JA, Howe PH. A regulated PNUTS mRNA to lncRNA splice switch mediates EMT and tumour progression. *Nature Cell Biology* 2017;**19**:1105-+.

39. Min KW, Davila S, Zealy RW, Lloyd LT, Lee IY, Lee R, Roh KH, Jung A, Jemielity J, Choi EJ, Chang JH, Yoon JH. eIF4E phosphorylation by MST1 reduces translation of a subset of mRNAs, but increases lncRNA translation. *Biochim Biophys Acta* 2017;**1860**:761-772.

40. Xue B, Leyrat C, Grimes JM, Robinson RC. Structural basis of thymosin-β4/profilin exchange leading to actin filament polymerization. *Proceedings of the National Academy of Sciences* 2014;**111**:E4596-E4605.

41. Yarmola EG, Parikh S, Bubb MR. Formation and implications of a ternary complex of profilin, thymosin β4, and actin. *Journal of Biological Chemistry* 2001;**276**:45555-45563.

42. Al Haj A, Mazur AJ, Buchmeier S, App C, Theiss C, Silvan U, Schoenenberger CA, Jockusch BM, Hannappel E, Weeds AG. Thymosin beta4 inhibits ADF/cofilin stimulated F- actin cycling and hela cell migration: Reversal by active Arp2/3 complex. *Cytoskeleton* 2014;**71**:95-107.

43. Mannherz HG, Hannappel E. The β- thymosins: intracellular and extracellular activities of a versatile actin binding protein family. *Cell motility and the cytoskeleton* 2009;**66**:839-851.

44. Mellad JA, Warren DT, Shanahan CM. Nesprins LINC the nucleus and cytoskeleton. *Current opinion in cell biology* 2011;**23**:47-54.

45. Zhou C, Rao L, Shanahan CM, Zhang Q. Nesprin-1/2: roles in nuclear envelope organisation, myogenesis and muscle disease. *Biochemical Society Transactions* 2018;**46**:311-320.

46. King SJ, Nowak K, Suryavanshi N, Holt I, Shanahan CM, Ridley AJ. Nesprin- 1 and nesprin-2 regulate endothelial cell shape and migration. *Cytoskeleton* 2014;**71**:423-434.

47. Revenu C, Athman R, Robine S, Louvard D. The co-workers of actin filaments: from cell structures to signals. *Nature reviews Molecular cell biology* 2004;**5**:635.

(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 lympthatic 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.

13

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 cm^2], 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 cm^2], 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.

Figure(s)
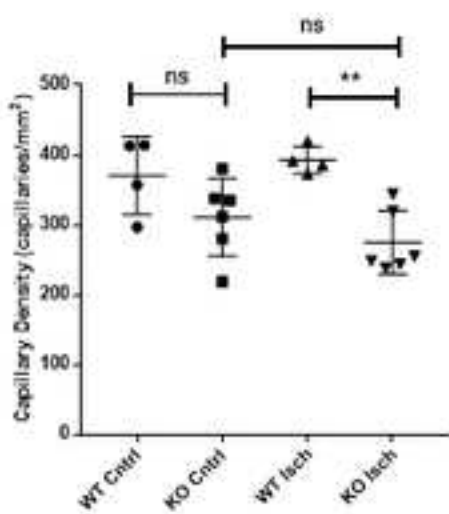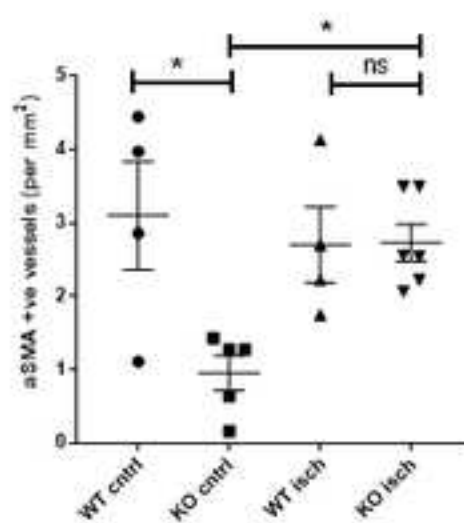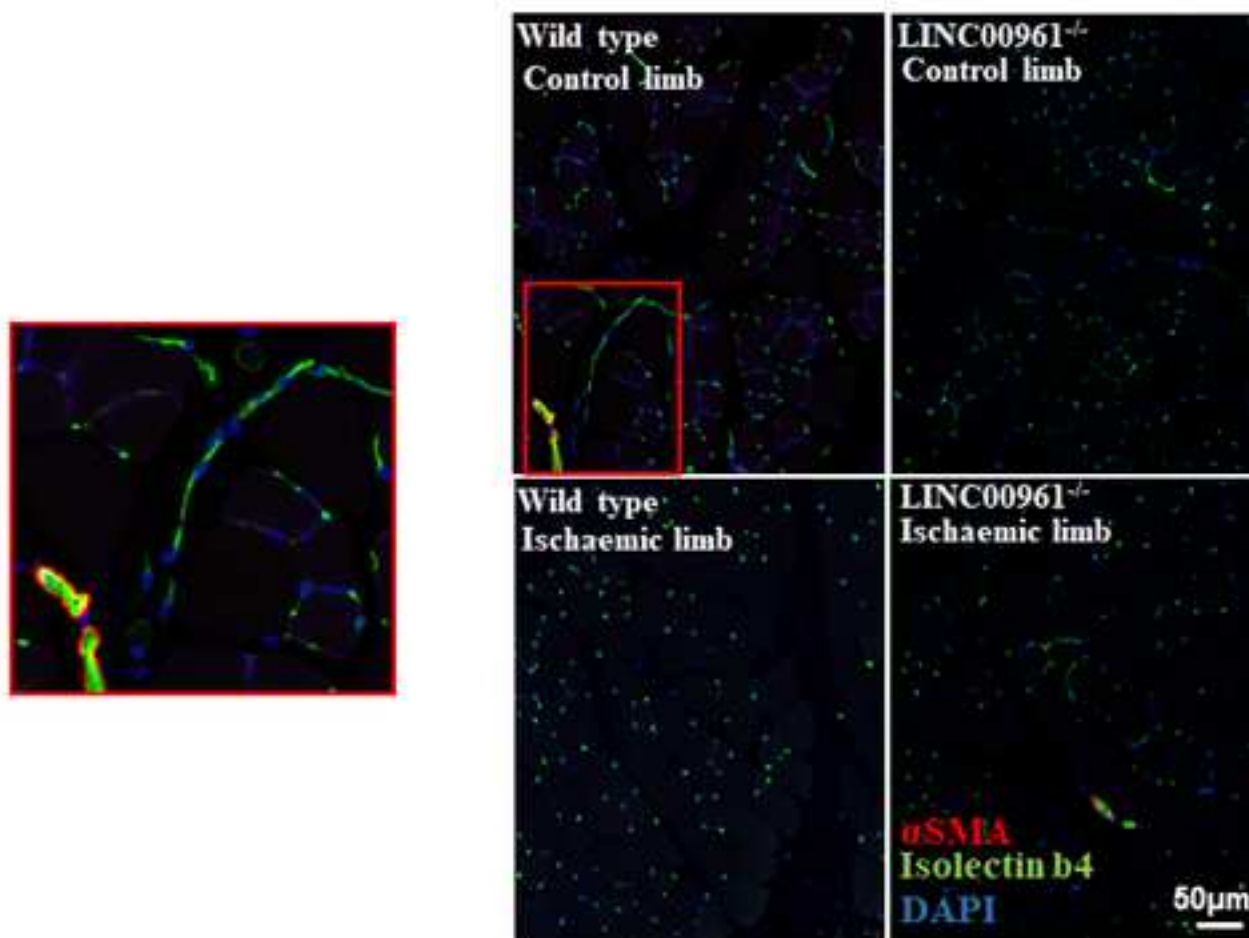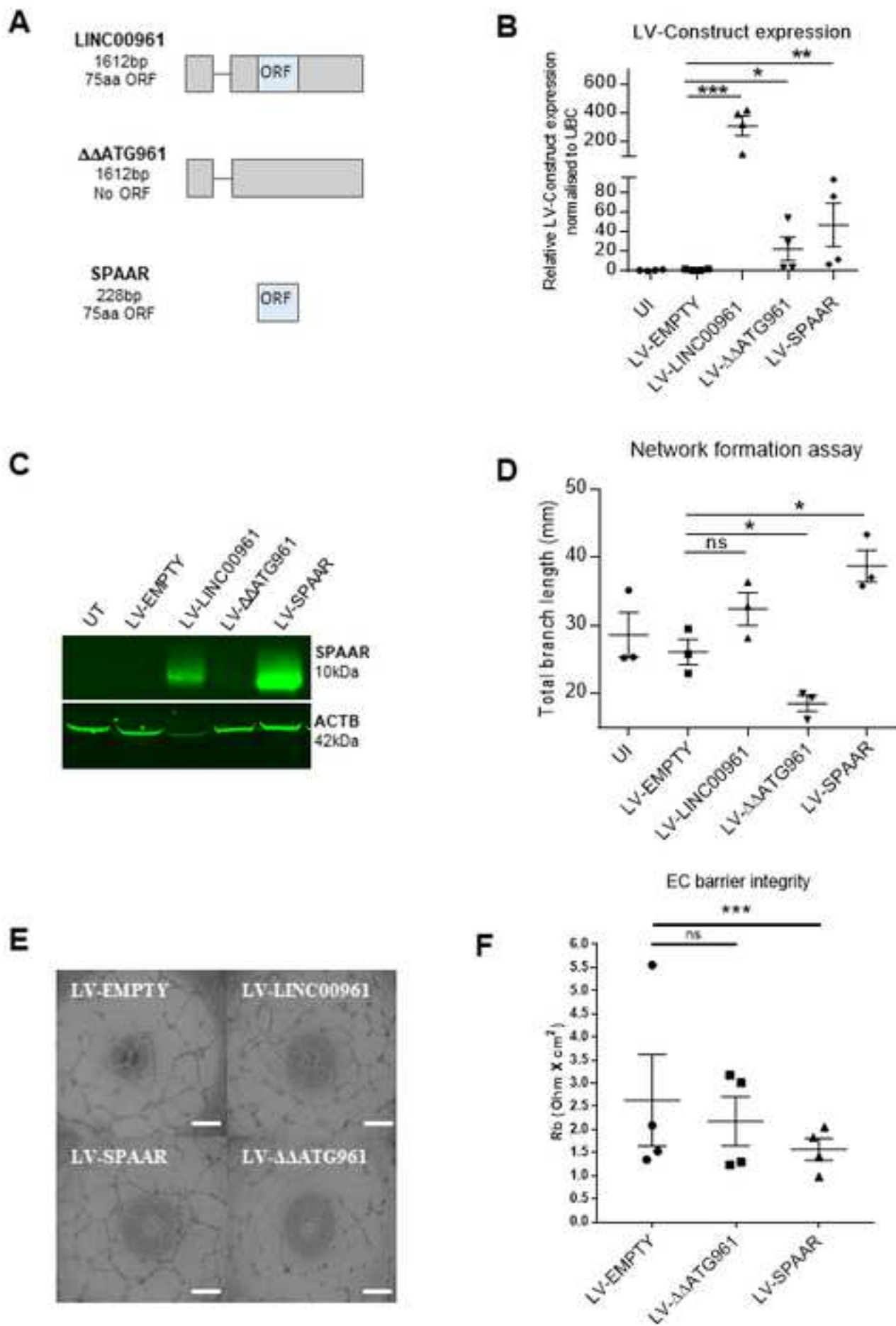
A LINC00961 expression

***

***

Relative LINC00961 expression normalised to UBC

1.5

1.0

0.5

0.0

Mock   Si Control   Si LINC00961

B Network formation assay

**

**

Total branch length (mm)

25

20

15

10

5

0

Mock   Si Control   Si LINC00961

C

Phase

Mock   Si Control   Si LINC00961

Calcein AM

Mock   Si Control   Si LINC00961

D Adhesion

**

**

Cells adhered (FC vs Mock)

1.5

1.0

0.5

0.0

Mock   Si Control   Si LINC00961

E EC Barrier Integrity

**

Rb (Ohm X cm^2)

3.5

3.0

2.5

2.0

1.5

1.0

0.5

0.0

Mock   SiControl   SiLINC00961

A

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

B

C

D

Wild type Control limb

LINC00961⁻/⁻ Control limb

Wild type Ischaemic limb

LINC00961⁻/⁻ Ischaemic limb

αSMA
Isolectin b4
DAPI

50μm

Figure(s)

Figure(s)

A

Network formation assay

B

Figure(s)

A

LINC00961
1612bp
75aa ORF

ORF

ΔΔATG961
1612bp
No ORF

SPAAR
228bp
75aa ORF

ORF

B

LV-Construct expression

Relative LV-Construct expression normalised to UBC

600
400
200
80
60
40
20
0

***
*
**

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

C

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

SPAAR
10kDa

ACTB
42kDa

D

Network formation assay

Total branch length (mm)

50
40
30
20

ns
*
*

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

E

LV-EMPTY   LV-LINC00961

LV-SPAAR   LV-ΔΔATG961

F

EC barrier integrity

Rb (Ohm X cm^2)

6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

ns
***

LV-EMPTY   LV-ΔΔATG961   LV-SPAAR

A

LINC00961 RNA pulldown

Biotinylated
LINC00961

HUVEC lysate

Proteins

Streptavidin
bead

SPAAR peptide pulldown

Lysate of HUVEC overexpressing HA-SPAAR

SPAAR HA

SPAAR HA Anti-HA
bead

Mass Spectrometry analysis

B

MARCKS
CALR
RPF2
MSI2
PRKCDBP
APEX1
FAM50A
EIF4H
DAZAP1
TMSB4X

0   5.0×10⁷   1.0×10⁸   1.5×10⁸
Label Free Quantification Intensity

C

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

0   1.0×10⁹   2.0×10⁹   3.0×10⁹
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

0   1   2   3
-Log10(pvalue)

E

GO Terms : SPAAR peptide pulldown

humoral immune response
leukocyte mediated immunity
proteolysis
response to axon injury

0.0   0.5   1.0   1.5   2.0
-Log10(pvalue)

A Network formation assay

B

Si Control   Si TMSB4X #1   Si TMSB4X #2

Si Control   Si TMSB4X #1   Si TMSB4X #2