American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
editorial@hematology.org

Single Cell Transcriptional Analysis of Human Endothelial Colony Forming Cells from Patients with Low VWF Levels

Tracking no: BLD-2021-010683R2

Christopher Ng (University of Colorado - Anschutz Medical Campus, United States) Alice Liu (University of Colorado - Anschutz Medical Campus, United States) Sujatha Venkataraman (University of Colorado - Anschutz Medical Campus, United States) Katrina Ashworth (Washington University School of Medicine, United States) Christopher Baker (University of Colorado - Anschutz Medical Campus, United States) Rebecca O'Rourke (University of Colorado - Anschutz Medical Campus, United States) Rajeev Vibhakar (University of Colorado - Anschutz Medical Campus, United States) Kenneth Jones (University of Oklahoma School of Medicine, United States) Jorge Di Paola (Washington University in St. Louis School of Medicine, United States)

Abstract:
Von Willebrand Factor (VWF) plays a key role in normal hemostasis and deficiencies of VWF lead to clinically significant bleeding. We sought to identify novel modifiers of VWF levels in endothelial colony forming cells (ECFCs) using single cell RNA sequencing (scRNA-seq). ECFCs were isolated from patients with Low VWF levels (plasma VWF antigen levels between 30-50 IU/dL) and from healthy controls. Human umbilical vein endothelial cells were used as an additional control cell line. Cells were characterized for their Weibel Palade body (WPB) content and VWF release. scRNA-seq of all cell lines was performed to evaluate for gene expression heterogeneity and for candidate modifiers of VWF regulation. Candidate modifiers identified by scRNA-seq were further characterized with siRNA experiments to evaluate for effects on VWF. We observed that ECFCs derived from patients with Low VWF demonstrated alterations in baseline WPB metrics and exhibit impaired VWF release. scRNA-seq analyses of these endothelial cells revealed overall decreased *VWF* transcription, mosaicism of *VWF* expression, and genes that are differentially expressed in Low VWF ECFCs and control endothelial cells (control ECs). A siRNA screen of potential VWF modifiers provided further evidence of regulatory candidates, and one such candidate, *FLI1*, alters the transcriptional activity of *VWF*. In conclusion, ECFCs from Low VWF individuals demonstrate alterations in their baseline VWF packaging and release as compared to control ECs. scRNA-seq revealed alterations in *VWF* transcription and siRNA screening identified multiple candidate regulators of VWF.

Conflict of interest: COI declared - see note

COI notes: Dr. Christopher Ng has received consultancy honorarium from CSL Behring and Takeda. Dr. Jorge Di Paola has received consultancy honorarium from CSL Behring.

Preprint server: No;

Author contributions and disclosures: C.J.N. designed research, performed research, contributed vital analytical tools, analyzed data, and wrote the paper. A.L. designed research, performed research, contributed vital analytical tools, analyzed data, and wrote the paper. S.V. designed research, performed research, contributed vital new reagents, and wrote the paper. K.A. designed research, performed research, contributed vital analytical tools, analyzed data, and wrote the paper. C.D.B. designed research, contributed vital new reagents, and wrote the paper. R.O'R. performed research, contributed vital new analytical tools, analyzed data, and wrote the paper. R.V. contributed vital new reagents, analyzed data, and wrote the paper. K.L.J. designed research, performed research, contributed vital new analytical tools, analyzed data, and wrote the paper. J.D.P. designed research, analyzed data, and wrote the paper.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: Data will be uploaded into the NCBI dBGaP system and relevant dbGaP study ID/accession numbers provided. Until that time that they are available in dbGaP, raw data files will be provided upon request by communication with Dr. Christopher Ng (Christopher.ng@cuanschutz.edu). Per the University of Colorado, use of the data is limited to investigations regarding "blood disorders" and for nonprofit use.

Clinical trial registration information (if any):

Single Cell Transcriptional Analysis of Human Endothelial Colony Forming Cells from Patients with Low VWF Levels

Christopher J Ng[1]*, Alice Liu[1], Sujatha Venkataraman[1], Katrina Ashworth[2], Christopher D. Baker[1], Rebecca O'Rourke[1], Rajeev Vibhakar[1], Kenneth L. Jones[3] and Jorge Di Paola[2]*

[1]Department of Pediatrics, University of Colorado and Children's Hospital Colorado, Aurora, CO, USA; University of Colorado Anschutz Medical Campus, Aurora, CO, USA

[2]Division of Hematology Oncology, Department of Pediatrics, Washington University School of Medicine in St. Louis, St. Louis, MO, USA

[3]Department of Cell Biology, University of Oklahoma Health Sciences Center, Oklahoma City, OK. Harold Hamm Diabetes Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK

**Co-corresponding authors*

Running Title: Transcriptional Alterations in Low VWF ECFCs

Address correspondence to: Christopher J Ng, Department of Pediatrics, University of Colorado and Children's Hospital Colorado, MS 8302, 12800 E 19th Ave, Aurora, CO 80045; email: Christopher.ng@cuanschutz.edu and Jorge Di Paola Washington University in St. Louis, 660 S. Euclid Avenue, Campus Box 8208, 5th floor MPRB St. Louis, MO 63110; email: dipaolaj@wustl.edu

Total Word Count: 4385

Key Points

1. ECFCs of patients with Low VWF exhibit decreased VWF secretion, transcriptional heterogeneity, and alterations in *VWF* mRNA content.
2. Single cell RNA sequencing of ECFCs identified *FLI1* and other genes as candidate regulators of VWF.

Abstract

Von Willebrand Factor (VWF) plays a key role in normal hemostasis and deficiencies of VWF lead to clinically significant bleeding. We sought to identify novel modifiers of VWF levels in endothelial colony forming cells (ECFCs) using single cell RNA sequencing (scRNA-seq). ECFCs were isolated from patients with Low VWF levels (plasma VWF antigen levels between 30-50 IU/dL) and from healthy controls. Human umbilical vein endothelial cells were used as an additional control cell line. Cells were characterized for their Weibel Palade body (WPB) content and VWF release. scRNA-seq of all cell lines was performed to evaluate for gene expression heterogeneity and for candidate modifiers of VWF regulation. Candidate modifiers identified by scRNA-seq were further characterized with siRNA experiments to evaluate for effects on VWF. We observed that ECFCs derived from patients with Low VWF demonstrated alterations in baseline WPB metrics and exhibit impaired VWF release. scRNA-seq analyses of these endothelial cells revealed overall decreased *VWF* transcription, mosaicism of *VWF* expression, and genes that are differentially expressed in Low VWF ECFCs and control endothelial cells (control ECs). A siRNA screen of potential VWF modifiers provided further evidence of regulatory candidates, and one such candidate, *FLI1*, alters the transcriptional activity of *VWF.* In conclusion, ECFCs from Low VWF individuals demonstrate alterations in their baseline VWF packaging and release as compared to control ECs. scRNA-seq revealed alterations in *VWF* transcription and siRNA screening identified multiple candidate regulators of VWF.

Introduction

VWF is a key hemostatic plasma protein that mediates the adhesion of platelets to sites of vascular injury and contributes to primary hemostasis. Significant deficiency of VWF leads to the bleeding diathesis known as von Willebrand disease (VWD) and individuals with VWD have a propensity to mucocutaneous bleeding that can have significant effects on morbidity and mortality.[1] The diagnosis of VWD is based on the presence of mucocutaneous bleeding and plasma VWF levels below the normal range, established as < 50 IU/dL by most clinical laboratories. Individuals with VWF levels < 30 IU/dL often have mutations in the *VWF* gene.[2] However, the rate of *VWF* mutations decreases significantly when VWF levels are between 30-50 IU/dL, suggesting the presence of other modifiers or gene variants that regulate these levels. Individuals with VWF levels between 30-50 IU/dL and bleeding are generally classified as Low VWF patients.

The wide distribution of plasma VWF levels in healthy individuals suggest multiple regulators of VWF. Previous reports have identified *ABO* as a significant modifier of VWF levels.[3] More recently, single nucleotide polymorphism (SNPs) in genes such as *STXBP5, SCARA5, STAB2, STX2, TC2N* and *CLE4M* have been associated with changes in VWF antigen levels.[3–6] Many of these gene variants have been implicated in the variability of VWF levels either through VWF clearance or WPB exocytosis.[7,8]

The transcriptional regulation of *VWF* has been partially characterized with reports investigating the roles of specific transcriptional and epigenetic regulators of *VWF* expression such as GATA, ETS1/2, H1, NF1, NFY, NFAT5 and more recently, miR-24 and *VWF* promoter methylation.[9–14] However, most of these studies of *VWF* transcriptional activity rely on non-endothelial cell lines

and over-expression models, which have been useful for the study of mutations in *VWF* but are arguably difficult to interpret in the setting of an intact *VWF* gene in the relevant cell type.

While megakaryocytes synthesize platelet VWF, it has been recently established that the majority of plasma VWF is derived from the vascular endothelium,[15,16] underscoring endothelial cells as a relevant model to study alterations in VWF regulation. In recent years, endothelial cells derived from peripheral blood, known as endothelial colony forming cells (ECFCs),[17] have shown to be useful in the study of VWF and VWD by demonstrating defects in VWF processing and localization, consistent with the clinical diagnoses of the patients from whom these cells were initially isolated.[18,19] These reports have also shown impairments in VWF release and abnormalities in WPB size and number.[18,19] It has been previously shown that alterations in WPB size are associated with altered VWF function.[20]

In this report we use ECFCs from a set of individuals with Low VWF levels (plasma concentrations of VWF in the 30-50 IU/dL range) to further elucidate mechanisms of VWF regulation outside of *VWF.* We hypothesized that ECFCs from these individuals would reveal variation in the processing, storage, and release of VWF when compared to individuals with higher plasma VWF levels. Through an unbiased transcriptional analysis of these individual's ECFCs we reveal novel genes that may regulate VWF expression or release.

Material and Methods

Materials

For further references to materials/reagents please see supplemental methods.

Patient enrollment

Individuals with plasma VWF:Ag levels between 30-50 IU/dL (defined as Low VWF patients) at the time of study enrollment were enrolled in our IRB-approved study at the University of Colorado (COMIRB #15-1072). ECFCs were isolated from 5 individuals with Low VWF Levels. The two control ECFC lines were from individuals who had plasma VWF levels of 75 and 127 IU/dL, respectively. ECFCs were isolated from whole blood samples as previously published (see Supplemental Figure 1).[21] Demographics and laboratory characteristics for the patients and controls are shown in Supplemental Table 1. We conducted our study/investigation using human subjects who signed informed consents under an IRB-approved protocol here at the University of Colorado.

Cell lines

After ECFC outgrowth, cells were assayed for cell surface markers (CD31, CD34, CD45, CD105, CD133, CD146) by flow cytometry (Supplemental Figure 1). Respective colonies were expanded and stored in liquid nitrogen. For each experiment, cells were maintained in EBM2-MV (Lonza) with supplemented FBS to 10%. Human umbilical vein cells (HUVECs) were purchased from Lonza (Portsmouth, NH). All endothelial cells used in experiments described here were passaged less than 6 times and experiments were synchronized to ensure similar passages/expansions. For this report, 5 Low VWF ECFCs and 3 Control ECs (2 control ECFCs and 1 HUVEC cell line) were used for all experiments unless otherwise noted.

VWF sequencing

VWF was sequenced as previously reported.[22] Details are shown in supplemental methods.

VWF release assays

VWF release assays were conducted as previously reported.[23] In brief, confluent endothelial cells were incubated in serum free media for a minimum of 2 hours before 1 hour of incubation with 25 or 50ng/mL of PMA. After treatment, supernatant and lysate VWF levels were assessed using ELISA assays as previously described.[24]

WPB characterization

Low VWF ECFCs and Control ECs were plated on collagen coated coverslips (Neuvitro, Vancouver, WA) at cellular density of 7,500cells/cm^2 and incubated for 24 hours. Cells were then fixed with paraformaldehyde (4%) and then permeabilized with Triton-X prior to incubation with Alexa 488 labeled anti-VWF antibody (Serotec, Hercules, CA). Coverslips were mounted onto glass slides and images acquired via an Olympus (Center Valley, PA) IX83 inverted microscope with CellSens software. A minimum of 9 images at pre-determined locations (to minimize bias) were captured. Images were analyzed in FIJI for the following WPB characteristics: total number, average size, and average aspect ratio via an automated image analysis macro (source code available upon request).

Single cell RNA sequencing analysis (scRNA-seq)

To interrogate transcriptional regulation and identify candidate regulators of VWF levels, we next analyzed the ECFC lines by performing scRNA-seq. Low VWF ECFCs were compared against three control endothelial cell lines: ECFC3, ECFC24, and HUVECs. HUVECs were used as their transcriptional pattern has been previous reported and they have been well characterized.[25] We sequenced a total of 10,481 cells from the three control and five Low VWF

ECFC cell lines. Endothelial cells were analyzed for scRNA-seq via the 10X Genomics platform and Illumina Novaseq 6000 platforms. We sequenced approximately 1,000 cells/sample with a read depth of approximately 50,000 reads/cell. Read mapping, and expression quantification was performed using a combination of the 10X Cellranger pipeline and custom analytic scripts. RNA copy number amounts are reported in unique molecular identifiers, which is the standard of the 10X Genomics analysis.[26] Briefly, single-cell reads were mapped to the human genome (GRCh38) and assigned to genes using the standard CellRanger pipeline. For the scRNA-seq analysis, the R packages Monocle and Seurat were used for differential expression and trajectory mapping (source code available upon request).[27–29] For the histogram analysis, cells were analyzed in R using the normalMixEM package (source code available upon request).[30] Heatmaps were generated using the Morpheus Software Package (https://software.broadinstitute.org/morpheus). In a downstream analysis, differentially expressed genes were analyzed using pathway analysis (Ingenuity Pathway Analysis).

siRNA screen

HUVEC supernatants and lysates were analyzed for VWF content after siRNA transfection with candidate genes. See supplemental methods for further details.

***VWF* promoter-GFP siRNA assay**

A lentiviral construct with a 1270bp segment of the *VWF* promoter 1270bp (-1023 to +247) was purchased from VectorBuilder (Chicago, IL) and transduced into HUVECs, forming eGFP-HUVECs. The eGFP-HUVECs were transfected with siRNAs as above for 72 hours and eGFP expression was determined using a Synergy H2 plate reader.

qPCR assays

See supplemental methods for further details.

Statistical analysis

All experiments were conducted with technical replication ($n \geq 3$) except for scRNA-seq. See supplemental methods for further details.

Data Sharing Statement

Data will be uploaded into the NCBI dBGaP system and relevant dbGaP study ID/accession numbers provided. Until that time that they are available in dbGaP, raw data files will be provided upon request by communication with Dr. Christopher Ng (Christopher.ng@cuanschutz.edu). Per the University of Colorado, use of the data is limited to investigations regarding “blood disorders” and for nonprofit use.

Results

VWF sequence variants

We sequenced *VWF* in all cell lines to interrogate potential genetic variants affecting VWF levels. Low VWF ECFC12 demonstrated a likely splice site variant (c.3108+1G>T), Low VWF ECFC17 demonstrated the p.Y1584C variant which has been associated with mild VWD type 1,[31] and Low VWF ECFC28 demonstrated both an intron variant and an exon 30 variant that have not been reported in gnomAD or ClinVar making the interpretation for pathogenicity difficult (Supplemental Table 1). Both Control ECFC24 and Low VWF ECFC12 also demonstrated the common variant p.D1472H which has been associated with abnormal Ristocetin cofactor activity but not with alterations in VWF:Ag levels or bleeding.[32]

Low VWF ECFCs have decreased number of WPBs and exhibit impaired VWF release

To determine if ECFCs demonstrated alterations in VWF content or packaging we analyzed WPB size and shape as well as constitutive and stimulated VWF release. Immunofluorescent image analysis of Low VWF ECFCs demonstrated a modest but statistically significant decrease in WPB quantity, average WPB size, and WPB aspect ratio as compared to control ECs (Supplemental Figure 2). Regarding VWF release, VWF levels in the cellular supernatant were not significantly different between control ECs and Low VWF ECFCs in the absence of PMA. However, when ECFCs were stimulated with 50ng/mL of PMA, there was a significant impairment in VWF release in the Low VWF ECFCs when compared to control ECs (Supplemental Figure 2). Although not statistically significant, there was also a trend towards increasing VWF amounts in the cellular lysates in the Low VWF group.

Transcriptional profile of ECFCs

A principal component analysis (PCA) comparing our sequenced endothelial cells to publicly available endothelial RNA-sequencing datasets demonstrates that our ECFCs and HUVECs cluster in proximity indicating that they were transcriptionally similar (Supplemental Figure 3). Additionally, high gene expression of venous markers (*NRP2, EPHB4)* and low gene expression of arterial markers (*EFNB2, NOTCH1)* suggest that our ECFCs were of a venous phenotype (Supplemental Figure 4). We defined a cell as endothelial if the cells were positive for any of *CDH5*, *PECAM1, ROBO4, ESAM, TIE1, or NOTCH4* based on a multi-tissue evaluation of highly expressed endothelial transcripts.[33] In screening with this methodology, 10,186 of 10,481 (97.1%) total cells were deemed to be endothelial (Figure 1A). There was no difference in the number of endothelial vs non-endothelial cells in each particular cell line (Figure 1B). After removing non-endothelial cells, we generated a t-distributed stochastic neighbor embedding (TSNE) plot that displayed the cells by their respective cell lines which reveals that each cell line largely occupies its own cluster location (i.e., difference between each individual cell line) but also displays transcriptional heterogeneity among the cells of a cell line (Figure 1C). When

identified by their Low VWF ECFC vs. control EC phenotype, there did not appear to be a significant clustering of cells by Low VWF vs control EC phenotype (Figure 1D).

Global transcriptional analysis identifies multiple pathways and genes that are differentially expressed in Low VWF ECFCs

We next determined differentially expressed genes between Low VWF ECFCs and Control ECs. There were 5,551 statistically significant differentially expressed genes ($p < 0.05$) when comparing control ECs to Low VWF ECFCs (Figure 1E). To ensure that these genes were specific to the comparison of Low VWF ECFCs vs. Control ECs and not driven by random variation between cell lines, we compared the top 500 of these genes vs genes differentially expressed in an ANOVA analysis of the 3 control cell lines and demonstrated little overlap between these two analyses (Supplemental Figure 5). Ingenuity Pathway Analysis demonstrated that Low VWF ECFCs showed statistically significant upregulation of specific pathways such as Integrin Signaling and Oxidative Phosphorylation (Figure 1F). Of note was the number of pathways implicated in our analysis that were associated with angiogenesis, such as Hypoxia Signaling in the Cardiovascular System and Rac Signaling.[34,35] A heatmap analysis reveals that the control endothelial cells (HUVEC, ECFC3, and ECFC24) cluster together by unbiased hierarchical clustering (Figure 1G).

scRNA-seq reveals transcriptional downregulation of *VWF* mRNA levels and *VWF* expression heterogeneity

As the Low VWF ECFCs are derived from individuals with decreased plasma VWF antigen levels, we next sought to investigate whether these levels were associated with decreased *VWF* mRNA expression. scRNA-seq reveals a significant decrease in *VWF* mRNA transcripts in the Low VWF ECFCs when compared to control ECFCs, (5.341 vs 9.076 unique molecular identifiers [UMI]/cell, $p<0.0001$). Although there was a global decreased in mRNA transcripts in

Low VWF ECFCs, this was mostly driven by ECFCs 12, 14, 17 and 28, as ECFC11 showed comparable *VWF* mRNA transcripts to control ECFCs (Figure 2A). Having noted this difference in *VWF* expression, we next determined if there was endothelial heterogeneity or mosaicism in *VWF* expression as previously identified by other groups.[13] Using all analyzed endothelial cells, we defined three populations of *VWF*-expressing cells, "high", "low" and "no expression." "High *VWF*" expressing cells had *VWF* UMIs higher than the average *VWF* expression (7 UMI/cell); these cells made up 39% of the population (Figure 2B). "Low *VWF*" cells had *VWF* UMIs equal to or lower than the average and they constituted the majority of cells (56%) (Figure 2B). "No *VWF*" expressing cells had no *VWF* expression (6%). We next examined the distribution of *VWF* expression in control ECs and Low VWF ECFCs using Gaussian mixed model fitting of *VWF* expression.[36] This analysis revealed multiple algorithmically-defined subpopulations of *VWF*-expressing cells (Figure 2C/2D). Interestingly, in control cells there appeared to be a bi-modal expression pattern with two peaks, the first in the 0-1 range and then the second in the 8-10 range. However, in our Low VWF ECFCs, there was a left-shift of *VWF* expression with a loss of the second peak of *VWF* expression. This finding was further supported by an evaluation of the "High *VWF*", "Low *VWF*," and "No *VWF*" populations in each cell line (Supplemental Figure 6), which demonstrates that Low VWF ECFCs generally have higher percentages of "'Low *VWF*" and "No *VWF*" cells. These data indicate that in our Low VWF ECFCs there was a shift towards a decrease in overall *VWF* transcription. Having noted the heterogeneity in *VWF* expression in both Control ECs and Low VWF ECFCs, we next hypothesized that this heterogeneity could be due to global changes in the endothelial gene transcripts, perhaps due to asynchronous cellular states. To analyze the potential for these processes, we utilized the "pseudotime" feature of the Monocle R package. This analysis revealed no significant differences between Control ECs and Low VWF ECFCs (Supplemental Figure 7).

Cluster-based analysis of differential gene expression

Having identified differences in *VWF* expression between Low VWF ECFCs and control ECs, we hypothesized that we could identify specific subsets of endothelial cells that would display the strongest differential expression of *VWF*. All ECs were analyzed via a TSNE plot to define clusters of ECs irrespective of Control EC or Low VWF ECFC status (Figure 3A). Each cluster was then analyzed for the *VWF* and *CDH5* expression of the Control EC and Low VWF ECFCs that comprised the cluster (Figure 3D). This analysis revealed that four primary clusters, (4,5,6, and 8) appear to demonstrate decreased *VWF* expression in Low VWF ECFCs (Figure 3B). These changes in *VWF* expression were independent of changes in overall *CDH5* expression (Figure 3C), similar to that seen in our pseudotime analysis (Supplemental Figure 7). These results suggest that a subset of cells may drive the differential *VWF* expression between our cohorts and these changes were likely not part of a global change in canonical endothelial gene expression.

siRNA-based functional screening identifies candidate VWF regulatory genes

We initially identified 5,551 differentially expressed genes when comparing control ECs and Low VWF ECFCs. To refine this initial list, we performed two additional analyses to better identify candidate genes. First, we identified genes that were differentially expressed between high-*VWF* expressing cells (>7 UMI/cell) and low-*VWF* expressing cells (≤7 UMI/cell). This approach assesses candidate genes for their association with *VWF* expression regardless of subject/cellular phenotype. Next, we identified genes that were differentially expressed between high-*VWF* expressing cells and low-*VWF* expressing cells in control ECs to minimize any potential contribution of Low VWF ECFC variability. These three gene lists, named subject level-expression analysis (Control EC vs. Low VWF ECFC), all cells-expression analysis (High- vs. Low-*VWF* expression in all cells), and Control EC–expression analysis (High- vs. Low-*VWF* expression in only control cells) were combined. The summary of these three differential gene expression lists are shown in a circos plot, where purple lines show gene overlap between the

three analyses (Figure 4A). The circos plot shows strong overlap between the candidate genes generated by the three differential expression analyses (Figure 4A). The initial candidate genes were chosen based on being (1) transcription factors previously established to bind to the region of the *VWF* promoter and/or associated with endothelial biology (*FOXO3*, *FLI1*, *TCF4*),[37–39] (2) genes with previous or purported associations with VWF levels (*COL4A1, SCARA3, ANGPT2, POU2F2, VAMP3),*[3,40] or (3) genes that showed strong differential expression (*ADGRA2, ADGRG6, SP1, C10orf10)* in our scRNA-seq experiments. These genes were then analyzed via an siRNA knockdown assay assessing VWF protein levels in cellular supernatant and lysates (Figure 4B). Genes that showed statistical difference in VWF expression in the lysate and supernatant after siRNA knockdown included *SCARA3*, *ANGTP2*, *FOXO3*, *FLI1*, *TCF4*, *TSTD1* and *ADGRA2* (Figure 4B).

Validation of *FLI1* as a candidate regulator of *VWF* expression

We further evaluated *FLI1* (Friend leukemia integration 1 transcription factor) as the ETS transcription factors (which encompass *FLI1*) have been previously shown to have regulatory control over vascular differentiation and regulatory effects on *VWF.*[41] *FLI1* was differentially expressed in subject-level analysis (1.07 vs 0.67 UMIs/cell, Control EC vs Low VWF ECFC, respectively, p = 1.85e-94) (Figure 5A), the all cells-expression analysis (1.00 vs 0.72 UMIs/cell, High- vs. Low-*VWF* expression, respectively, p = 1.05e-42) (Figure 5B), and the control EC-expression analysis (1.17 vs 0.97 UMIs/cell, High- vs. Low-*VWF* expression, respectively, p = 2.94e-09) (Figure 5C). Knockdown of *FLI1* was confirmed via mRNA levels (Figure 5D) and western blot in siRNA-treated HUVECs (Supplemental Figure 8). *VWF* mRNA expression was decreased after *FLI1* knockdown (Figure 5E), supporting the possibility that *FLI1* may regulate *VWF* expression. A VWFpromoter-GFP reporter assay demonstrated decreased *VWF* promoter activity after *FLI1* knockdown as compared to control siRNA knockdown (Figure 5F) suggesting *FLI1* regulates *VWF* transcription.

Discussion

In this study we report a transcriptional analysis of ECFCs from a cohort of individuals with Low VWF levels. A recent report suggests that endothelial cells share a select number of genes with high correlation.[33] Using this gene signature, we demonstrate that our ECFCs exhibit an endothelial phenotype. While there have been concerns of whether ECFCs represent true endothelial cells, our comparative analysis to ENCODE RNA-seq data suggest that our ECFCs are transcriptionally similar to HUVECs. Regarding tissue specificity, our ECFCs seem most similar to venous or capillary (heart, lung, skeletal muscle, brain) endothelium and are less similar to bone marrow, kidney, or liver endothelium based on *VWF* expression patterns.[13,42–44] A previous study suggested that ECFCs had a transcriptional pattern similar to microvascular cells.[45] The discrepancy with our results is likely due to the fact that the study was based on microarray data with limited number of transcripts, and not on scRNA-seq that represents the whole transcriptome as we used in our report.

In our studies we demonstrate that ECFCs isolated from individuals with Low VWF levels display alterations in WPB size/shape and have decreased VWF release. These findings are similar to previous findings in ECFCs from patients with VWD and mutations in *VWF.*[18,19] However, in contrast to previous reports,[19] we found no significant decrease in ECFC VWF content. This difference may be explained by the more significant decrease in plasma VWF levels in the patients studied in Starke et al. as compared to our study.[19] Regarding our WPB findings, the decreased VWF release/secretory defect that we observed here is consistent with previously reported data,[19] and suggests that decreased VWF release may be associated with the Low VWF Level phenotype. Interestingly, our scRNA-seq analyses show that some genes associated with WPB-exocytosis[46,47] (*STXBP5*, *RAB3A*, *RAB15*, *MYRIP*, *MYO5A*, *UNC13D*) are differentially expressed in Low VWF ECFCs. Furthermore, RhoA and RAC signaling, both of which are associated with WPB exocytosis signaling,[48,49] are implicated in our IPA analysis.

Further work into WPB-specific genes and pathways in Low VWF ECFCs are warranted to better delineate the effects on WPB biology.

Our study also showed decreased *VWF* mRNA levels in Low VWF ECFCs, similar to previous reports in Type 1 VWD ECFCs.[19] The underlying molecular mechanism for the decrease in *VWF* mRNA levels is unclear; potential explanations include upstream transcriptional regulators or other aspects of RNA modulation, such as RNA stability, RNA splicing, and mRNA decay. There is evidence that suggests that RNA splicing may affect *VWF* expression and contribute to the pathophysiology of Low VWF levels and one subject has a potential splice site abnormality identified.[50,51] One limitation of our study is that we do not have comprehensive sequencing analysis of the *VWF* promoter and it has been reported that polymorphisms in the promoter could have an effect on *VWF* mRNA levels.[52,53]

Our TSNE analysis demonstrated transcriptional differences between cell lines and also differences within a single cell line. Consistent with the report by Yuan et al.,[13] we identified heterogeneous *VWF* expression in endothelial cells. Using similar mixed model analysis to Yuan et al., we demonstrated that there are potentially multiple *VWF* expressing subpopulations in our ECFCs. However, we also showed that the pattern of *VWF* expression is different between control ECs and Low VWF ECFCs, even though we did not find a clear differential clustering of Low VWF ECFCs vs control ECs. As shown in Figure 2C, Low VWF ECFCs are left-shifted when compared to control ECs with an overall reduction in *VWF* transcripts of approximately 40%. Interestingly, ECFC11 did not show a significant decrease in *VWF* mRNA content, raising the possibility that the low VWF levels observed in the patient where this cell line came from are caused by a different mechanism, perhaps via increased VWF clearance. Furthermore, our findings that a subset (clusters 4,5,6, and 8) of endothelial cells may be responsible for the differential *VWF* expression between control ECs and Low VWF ECFCs

further support the theory of heterogenous *VWF* expression in endothelial cells. Additional studies are required to evaluate these clusters to determine if they represent a definable subset of endothelial cells that drive *VWF* expression. While our data suggest that there is little evidence of significant global transcriptional differences between Low VWF ECFCs and control ECs, there may be subsets of endothelial cells that display differential gene expression perhaps through epigenetic[13] or other alternative mechanisms.

Our differential gene analysis led us to the identification of potential candidate genes via a combinatorial approach to identify genes associated with a phenotype (ECFCs from Low VWF patients) as well as those associated with *VWF* expression (High- vs. Low-*VWF* expression). We further analyzed these candidate genes using an siRNA-based screen and demonstrated that knockdown of a putative *VWF* transcription factor, *FLI1*, significantly decreases *VWF* at both the mRNA and protein levels, likely through reduced *VWF* promoter activity. Our results correlate with previous reports that showed a critical role of *FLI1* in vascular differentiation and that *FLI1* overexpression increased *VWF* promoter activity.[38,54,55] Our approach to screening and validation (as we did with *FLI1)* is similar to a recent GWAS analysis which utilized genomic (as opposed to transcriptomic) data to identify candidate regulatory genes that were further evaluated using siRNA assays.[4] Interestingly, some candidate genes like *RAB5C, ARSA and SYNGR1* found in a recent GWAS study[4] for VWF levels were also found to be differentially expressed in our transcriptional analysis. While siRNA-based knockdowns are a standard approach in the screening of candidate genes, siRNA transfection itself may lead to transcriptional changes. Therefore, other functional screening techniques, such as by CRISPR-Cas9, may better refine candidate genes. This may be important in the context of the Low VWF Level phenotype, as this clinical phenotype is a relatively mild decrease in VWF plasma levels and significant gene knockdown may not fully represent the *in vivo* biological mechanism of the Low VWF phenotype. A limitation of our approach is that we evaluated only a subset of

candidate genes, and our work here should not suggest that there is a single unifying mechanism driving the Low VWF Level phenotype.

There are several other limitations to our work. We surveyed only a relatively small number of ECFC cell lines but designed analysis pathways to minimize sampling bias and take maximal advantage of our scRNA-seq. We attempted to synchronize the expansions and passages of our cells to minimize the possibility of potential variations in ECFC characteristics over time as demonstrated in Boer et al.[56] This report found that different ECFC "groups" have different VWF expression patterns at the protein level,[56] although it is not known if these differences also occurred at the mRNA level. While it is possible our cell lines may have represented different "groups," all our cell lines appeared phenotypically similar to the observed description of "Group 1" ECFCs and thus it is less likely that we included ECFCs that were similar to the "Group 2" or "Group 3" phenotype. Another potential concern is the identified genetic variants in our Low VWF ECFCs and how these may relate to our findings. One of these variants, Y1584C, has been associated with Low VWF/Type 1 VWD but was shown to have a loss of high molecular weight multimers, which would be unlikely to affect transcriptional regulation.[57] While it is possible that the identified variants may affect transcriptional regulation of *VWF* (such as the possible splice site variant found in ECFC12), we would generally expect the other variants to affect final protein translation (or protein function in the case of Y1584C) and thus our transcriptional findings might be less likely to be affected by these genetic variants. There were three variants (c.3108+1G>T, c.1946-10T>G and c.5281dup) not present in human genomic databases or listed in ClinVar making their potential relevance difficult to assess. Future work specifically investigating these variants and their contribution to the transcriptional findings found here would be warranted. Finally, our analysis presented here focuses on a transcriptional analysis of ECFCs from Low VWF individuals. While we have found evidence of transcriptional alterations of *VWF* in these ECFC cell lines, there are other alternative

mechanisms that can potentially explain the clinical phenotype of these patients, such as alterations in VWF processing, secretion, or clearance.

In summary, we analyzed ECFCs from individuals with Low VWF levels using scRNA-seq analysis. We identified a candidate list of potential regulators of *VWF* and demonstrate that transcriptional regulation at the single cell level may play a role in the complex pathophysiology of the Low VWF phenotype. We suggest that transcriptional regulation may be yet another mechanism contributing to the Low VWF phenotype.

Authorship

C.J.N. designed research, performed research, contributed vital analytical tools, analyzed data, and wrote the paper. A.L. designed research, performed research, contributed vital analytical tools, analyzed data, and wrote the paper. S.V. designed research, performed research, contributed vital new reagents, and wrote the paper. K.A. designed research, performed research, contributed vital analytical tools, analyzed data, and wrote the paper. C.D.B. designed research, contributed vital new reagents, and wrote the paper. R.O'R. performed research, contributed vital new analytical tools, analyzed data, and wrote the paper. R.V. contributed vital new reagents, analyzed data, and wrote the paper. K.L.J. designed research, performed research, contributed vital new analytical tools, analyzed data, and wrote the paper. J.D.P. designed research, analyzed data, and wrote the paper.

Sources of Funding

This work was supported by grants/support from the National Hemophilia Foundation (CJN), the Health Resources & Services Administration – Maternal & Child Health (5 H30 MC00008-20-00 to CJN/JDP), the National Institutes of Health (R01 HL120728 and R01 HL141794 to JDP, P01

HL144457 to CN/JDP) and the Postle Family Chair for Pediatric Cancer and Blood Disorders (JDP).

Disclosures

Dr. Christopher Ng has received consultancy honorarium from CSL Behring and Takeda. Dr. Jorge Di Paola has received consultancy honorarium from CSL Behring.

References

1. Ng C, Motto DG, Paola JD. Diagnostic approach to von Willebrand disease. *Blood*. 2015;125(13):2029 2037.

2. Lavin M, Aguila S, Schneppenheim S, et al. Novel Insights into the Clinical Phenotype and Pathophysiology Underlying Low VWF Levels: The Low Von Willebrand Factor in Ireland Cohort (LoVIC) Study. *Blood*. 2016;128(22):873.

3. Smith NL, Chen M-H, Dehghan A, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation*. 2010;121(12):1382 1392.

4. Sabater-Lleal M, Huffman JE, Vries PS de, et al. Genome-Wide Association Transethnic Meta-Analyses Identifies Novel Associations Regulating Coagulation Factor VIII and von Willebrand Factor Plasma Levels. *Circulation*. 2019;139(5):620 635.

5. Loon J van, Dehghan A, Weihong T, et al. Genome-wide association studies identify genetic loci for low von Willebrand factor levels. *European journal of human genetics : EJHG*. 2016;24(7):1035 1040.

6. Sanders YV, Bom JG, Isaacs A, et al. CLEC4M and STXBP5 gene variations contribute to von Willebrand factor level variation in von Willebrand disease. *J Thromb Haemost*. 2015;13(6):956–966.

7. Zhu Q, Yamakuchi M, Ture S, et al. Syntaxin-binding protein STXBP5 inhibits endothelial exocytosis and promotes platelet secretion. *J Clin Invest*. 2014;124(10):4503 4516.

8. Swystun LL, Lai JD, Notley C, et al. The endothelial cell receptor stabilin-2 regulates VWF-FVIII complex half-life and immunogenicity. *J Clin Invest*. 2018;128(9):4057–4073.

9. Liu J, Kanki Y, Okada Y, et al. A +220 GATA motif mediates basal but not endotoxin-repressible expression of the von Willebrand factor promoter in Hprt-targeted transgenic mice. *J Thromb Haemost*. 2009;7(8):1384 1392.

10. Yuan L, Nikolova-Krstevski V, Zhan Y, et al. Antiinflammatory effects of the ETS factor ERG in endothelial cells are mediated through transcriptional repression of the interleukin-8 gene. *Circulation research*. 2009;104(9):1049 1057.

11. Liu J, Yuan L, Molema G, et al. Vascular bed-specific regulation of the von Willebrand factor promoter in the heart and skeletal muscle. *Blood*. 2011;117(1):342 351.

12. Peng Y, Jahroudi N. The NFY transcription factor functions as a repressor and activator of the von Willebrand factor promoter. *Blood*. 2002;99(7):2408 2417.

13. Yuan L, Chan GC, Beeler D, et al. A role of stochastic phenotype switching in generating mosaic endothelial cell heterogeneity. *Nat Commun*. 2016;7(1):10160.

14. Xiang Y, Cheng J, Wang D, et al. Hyperglycemia repression of miR-24 coordinately upregulates endothelial cell expression and secretion of von Willebrand factor. *Blood*. 2015;125(22):3377 3387.

15. Jaffe EA, Hoyer LW, Nachman RL. Synthesis of Antihemophilic Factor Antigen by Cultured Human Endothelial Cells. *The Journal of clinical investigation*. 1973;52(11):2757 2764.

16. Everett LA, Cleuren ACA, Khoriaty RN, Ginsburg D. Murine coagulation factor VIII is synthesized in endothelial cells. *Blood*. 2014;123(24):3697–3705.

17. Lin Y, Weisdorf DJ, Solovey A, Hebbel RP. Origins of circulating endothelial cells and endothelial outgrowth from blood. *J Clin Invest*. 2000;105(1):71–77.

18. Wang J-W, Bouwens EAM, Pintao MC, et al. Analysis of the storage and secretion of von Willebrand factor in blood outgrowth endothelial cells derived from patients with von Willebrand disease. *Blood*. 2013;121(14):2762 2772.

19. Starke RD, Paschalaki KE, Dyer CEF, et al. Cellular and molecular basis of von Willebrand disease: studies on blood outgrowth endothelial cells. *Blood*. 2013;121(14):2773 2784.

20. Ferraro F, da SML, Grimes W, et al. Weibel-Palade body size modulates the adhesive activity of its von Willebrand Factor cargo in cultured endothelial cells. *Scientific reports*. 2016;6(1):32473.

21. Baker CD, Black CP, Ryan SL, Balasubramaniam V, Abman SH. Cord blood endothelial colony-forming cells from newborns with congenital diaphragmatic hernia. *The Journal of pediatrics*. 2013;163(3):905 907.

22. Lavin M, Aguila S, Schneppenheim S, et al. Novel insights into the clinical phenotype and pathophysiology underlying low VWF levels. *Blood*. 2017;130(21):2344–2353.

23. Ng CJ, McCrae KR, Ashworth K, et al. Effects of anti-β2GPI antibodies on VWF release from human umbilical vein endothelial cells and ADAMTS13 activity. *Res Pract Thrombosis Haemostasis*. 2018;2(2):380–389.

24. White-Adams TC, Ng CJ, Jacobi PM, Haberichter SL, Paola JAD. Mutations in the D′D3 region of VWF traditionally associated with type 1 VWD lead to quantitative and qualitative deficiencies of VWF. *Thromb Res*. 2016;145:112–118.

25. Consortium TEP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57.

26. Genomics 10X. How does Cell Ranger correct for amplification bias? https://kb.10xgenomics.com/hc/en-us/articles/115004037743-How-does-Cell-Ranger-correct-for-amplification-bias-. Accessed 2021-04-07. 2021;

27. Qiu X, Hill A, Packer J, et al. Single-cell mRNA quantification and differential analysis with Census. *Nat Methods*. 2017;14(3):309–315.

28. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014;32(4):381–386.

29. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411.

30. Benaglia T, Chauveau D, Hunter DR, Young D. mixtools : An R Package for Analyzing Finite Mixture Models. *J Stat Softw*. 2009;32(6):.

31. James PD, Notley C, Hegadorn C, et al. The mutational spectrum of type 1 von Willebrand disease: Results from a Canadian cohort study. *Blood*. 2007;109(1):145 154.

32. Flood VH, Friedman KD, Gill JC, et al. No increase in bleeding identified in type 1 VWD subjects with D1472H sequence variation. *Blood*. 2013;121(18):3742 3744.

33. Butler LM, Hallström BM, Fagerberg L, et al. Analysis of Body-wide Unfractionated Tissue Data to Identify a Core Human Endothelial Transcriptome. *Cell Syst*. 2016;3(3):287-301.e3.

34. Krock BL, Skuli N, Simon MC. Hypoxia-induced angiogenesis: good and evil. *Genes Cancer*. 2011;2(12):1117–33.

35. Fryer BH, Field J. Rho, Rac, Pak and angiogenesis: old roles and newly identified responsibilities in endothelial cells. *Cancer Lett*. 2005;229(1):13 23.

36. Benaglia T, Chauveau D, Hunter DR, Young D. mixtools: AnRPackage for Analyzing Finite Mixture Models. *J Stat Softw*. 2009;32(6):.

37. Schwachtgen JL, Janel N, Barek L, et al. Ets transcription factors bind and transactivate the core promoter of the von Willebrand factor gene. *Oncogene*. 1997;15(25):3091 3102.

38. Nagai N, Ohguchi H, Nakaki R, et al. Downregulation of ERG and FLI1 expression in endothelial cells triggers endothelial-to-mesenchymal transition. *Plos Genet*. 2018;14(11):e1007826.

39. Baratz KH, Tosakulwong N, Ryu E, et al. E2-2 protein and Fuchs's corneal dystrophy. *New England Journal of Medicine*. 2010;363(11):1016 1024.

40. Flood VH, Schlauderaff AC, Jacobi PM, et al. VWF Interaction With Type IV Collagen Is Mediated Through Critical VWF A1 Domain Residues. *Blood*. 2013;122(21):29–29.

41. Groeneveld DJ, Bekkum T, Dirven RJ, et al. Angiogenic characteristics of blood outgrowth endothelial cells from patients with von Willebrand disease. *J Thromb Haemost*. 2015;13(10):1854–1866.

42. Aird WC, Edelberg JM, Weiler-Guettler H, et al. Vascular Bed–specific Expression of an Endothelial Cell Gene Is Programmed by the Tissue Microenvironment. *J Cell Biology*. 1997;138(5):1117–1124.

43. Yamamoto K, Waard V de, Fearns C, Loskutoff DJ. Tissue distribution and regulation of murine von Willebrand factor gene expression in vivo. *Blood*. 1998;92(8):2791–801.

44. Fujisawa T, Tura-Ceide O, Hunter A, et al. Endothelial Progenitor Cells Do Not Originate From the Bone Marrow. *Circulation*. 2019;140(18):1524–1526.

45. Toshner M, Dunmore B, McKinney E, et al. Analyses of blood outgrowth endothelial cells reveal an endothelial HOX gene signature in human beings. *Lancet*. 2013;381:S108.

46. Nightingale T, Cutler D. The secretion of von Willebrand factor from endothelial cells; an increasingly complicated story. *J Thromb Haemost*. 2013;11 Suppl 1(Suppl. 1):192 201.

47. McCormack JJ, Silva ML da, Ferraro F, Patella F, Cutler DF. Weibel-Palade bodies at a glance. *J Cell Sci*. 2017;130(21):3611 3617.

48. Schillemans M, Karampini E, Kat M, Bierings R. Exocytosis of Weibel–Palade bodies: how to unpack a vascular emergency kit. *J Thromb Haemost*. 2019;17(1):6–18.

49. Klarenbach SW, Chipiuk A, Nelson RC, Hollenberg MD, Murray AG. Differential actions of PAR2 and PAR1 in stimulating human endothelial cell exocytosis and permeability: the role of Rho-GTPases. *Circulation research*. 2003;92(3):272 278.

50. Yuan L, Janes L, Beeler D, et al. Role of RNA splicing in mediating lineage-specific expression of the von Willebrand factor gene in the endothelium. *Blood*. 2013;121(21):4404–4412.

51. Hawke L, Bowman ML, Poon M-C, et al. Characterization of aberrant splicing of von Willebrand factor in von Willebrand disease: an underrecognized mechanism. *Blood*. 2016;128(4):584 593.

52. Mojiri A, Alavi P, Carrillo MAL, et al. Endothelial cells of different organs exhibit heterogeneity in von Willebrand factor expression in response to hypoxia. *Atherosclerosis*. 2019;282:1–10.

53. Morini MF, Giampietro C, Corada M, et al. VE-Cadherin-Mediated Epigenetic Regulation of Endothelial Gene Expression. *Circulation research*. 2017;122(2):231 245.

54. Toyama T, Asano Y, Miyagawa T, et al. The impact of transcription factor Fli1 deficiency on the regulation of angiogenesis. *Exp Dermatol*. 2017;26(10):912–918.

55. Asano Y, Stawski L, Hant F, et al. Endothelial Fli1 Deficiency Impairs Vascular Homeostasis A Role in Scleroderma Vasculopathy. *Am J Pathology*. 2010;176(4):1983–1998.

56. Boer S, Bowman M, Notley C, et al. Endothelial characteristics in healthy endothelial colony forming cells; generating a robust and valid ex vivo model for vascular disease. *J Thromb Haemost*. 2020;18(10):2721–2731.

57. Pruss CM, Golder M, Bryant A, et al. Pathologic mechanisms of type 1 VWD mutations R1205H and Y1584C through in vitro and in vivo mouse models. *Blood*. 2011;117(16):4358–66.

Figure Legends

Figure 1: Control ECs and Low VWF ECFCs display global transcriptional differences

Endothelial cells from eight cell lines (5 Low VWF and 3 control lines) were sequenced via scRNA-seq (10X Genomics/Illumina).

1A. After filtering and normalization, a TSNE plot was generated to determine cells that were of an endothelial vs non-endothelial nature. Endothelial cells (orange) were determined to be a cell that was positive for *CDH5*, *PECAM1*, *ROBO4*, *ESAM*, *TIE1*, or *NOTCH4*. Cells that were negative for any EC marker are shown in purple.

1B. Bar graph demonstrating the overall distribution of endothelial vs non-endothelial cells (as determined by any positivity for *CDH5*, *PECAM1*, *ROBO4*, *ESAM*, *TIE1*, or *NOTCH4*) in each cell line. Bar graph represents the raw value of the number of endothelial vs non-endothelial cells in each cell line.

1C. After exclusion of non-endothelial cells, a TSNE plot was regenerated to show clustering broken down by individual cell line, displayed to determine overall differences in transcriptional profiles. Each cell line is displayed in a different color.

1D. After exclusion of non-endothelial cells, a TSNE plot was regenerated with endothelial cells highlighted as either Control EC (red) or Low VWF ECFC (blue) to determine transcriptional differences between the two populations.

1E. Volcano plot demonstrating differential expression between control ECs and Low VWF ECFCs. The y-axis is the $\log_{10}$ of the P-value and the x-axis is the $-\log_2$ value of fold change. Each individual dot represents a statistically significant differential expressed gene ($P<0.05$) as measured by P-value adjusted for multiple comparisons. Genes shown in blue are more highly expressed in control ECs as compared to Low VWF ECFCs and genes shown in red are more highly expressed in Low VWF ECFCs as compared to control ECs.

1F: Ingenuity pathway analysis (IPA) showing the most highly implicated pathways between the control ECs and Low VWF ECFCs. Pathways are ordered by the statistical strength of the activation z-score (which measures how differentially regulated a pathway is between the two cohorts). A positive z-score (as denoted by the darker blue on the scale) suggest that the pathway is activated in the Low VWF ECFCs as compared to control ECs. For some pathways, shown in gray, IPA was unable to predict a directionality to the change in regulation (activation or inhibition) but there was a high degree of differentially expressed genes in that particular pathway.

1G. Heatmap of the top 250 most differentially expressed genes between control ECs and Low VWF ECFCs demonstrating hierarchical clustering of the three control cell lines closely together.

Figure 2: Evidence of transcriptional downregulation and left-shifting of *VWF* expression

2A. The average *VWF* expression demonstrates a significant decrease in *VWF* mRNA copy number in the Low VWF ECFCs as compared to control ECs. The average VWF expression for all cells was 7 UMI/cell.

2B. A waffle plot of all sequenced cells (Low VWF ECFCs and Control ECs) demonstrating the overall distribution of High (>7 UMIs of *VWF*/cell) *VWF* expression, Low *VWF* Expression (≤7 UMIs of *VWF*/cell), and no *VWF* expression (0 UMIs of *VWF*/cell).

2C/D. Normalized histogram plots of *VWF* expression (UMI) for the control EC (**2C**) and Low VWF ECFCs (**2D**) cohorts. A mixed modeling algorithm determines the number of gaussian populations that may make up the overall distribution of the histogram. 4 sub-populations (identified by the pink, teal, purple, and green curves), each with their own gaussian distribution, are predicted to create the overall population distribution in both the control ECs and the Low VWF ECFCs. Statistical significance analyzed with Mann-Whitney U test, significance is shown with asterisks (**** $p < 0.0001$).

Figure 3: Cluster-based differential gene expression of *VWF* and *CDH5*

3A. TSNE plot demonstrates 8 primary clusters of endothelial cells as determined by their overall transcriptional pattern, regardless of initial control EC vs Low VWF ECFC phenotype.

3B/3C. Each cluster identified in **3A** is then analyzed for *VWF* expression and *CDH5* expression as shown in **3B** and **3C**, respectively, with the top row of each gene in **3B** and **3C** corresponding to clusters 1-4, and the bottom row corresponding to rows 5-8.

3D. Finally, the relative distribution (% of total) of control ECs vs Low VWF ECFCs are shown for each individual cluster.

Figure 4: Composite candidate genes and siRNA-based screen of VWF supernatant and lysate levels after target gene knockdown

4A. Circos plot demonstrating overlap of candidate genes from the most highly differentially expressed genes for three independent differential expression analyses: (1) Control EC vs Low VWF ECFCs (subject level-expression analysis), (2) All cells-expression analysis and (3) Control EC-expression analysis. Genes are visually shown as composing the shaded outer

circle (green – differentially expressed genes in subject level-expression analysis, red – differentially expressed genes in the All cells-expression analysis, and blue – differentially expressed genes in the Control EC-expression analysis). The dark orange shading on the inner circle represents genes that are represented in more than 1 analyses and are connected to overlapping genes in another analysis via purple lines. Light orange shading represents genes that are unique to only that differential expression analysis.

4B. Candidate genes were assessed to determine effects on VWF protein levels in HUVECs via an siRNA-knockdown assay. After transfection with 1nM of candidate siRNAs (2-4 siRNAs per gene) VWF protein levels were assessed in the cellular lysates and cellular supernatants via a VWF ELISA. Values are reported as % expression of control scrambled siRNA. For all experiments, N>3 and p-values of significant relationships are shown as analyzed by a one-way ANOVA with Holm-Sidak's multiple comparison test. Error bars represent the means ± SEM. Statistical significance is shown with asterisks as compared to control siRNA (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p < 0.0001$).

Figure 5: *FLI1* is a candidate regulator for VWF based on differential expression and siRNA knockdown

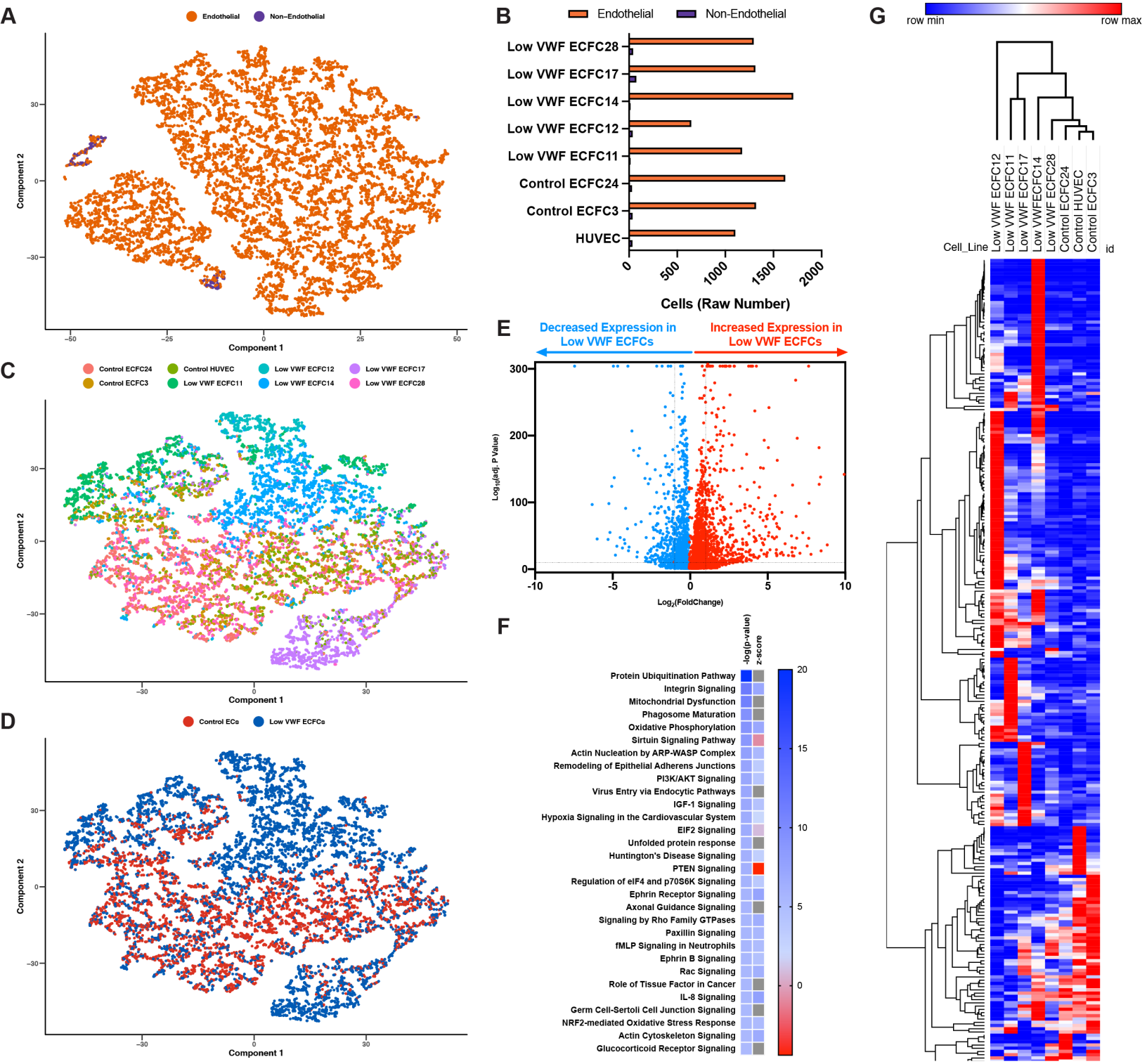
5A/5B/5C. scRNA-seq data demonstrated decreases in *FLI1* expression in three expression analyses, subject level-expression analysis (**5A**), All cells-expression analysis (**5B**), and Control EC–expression analysis (**5C**).

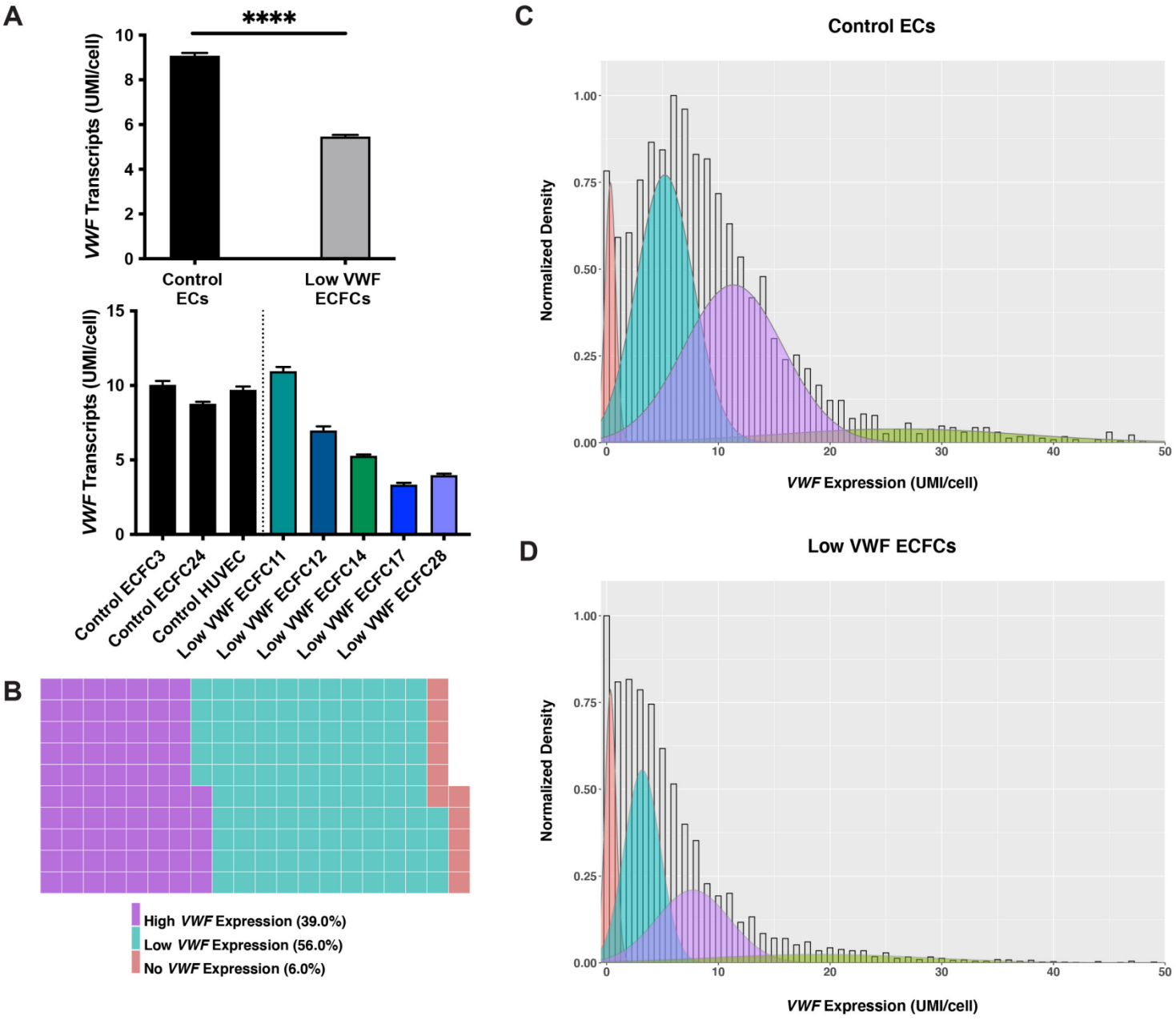
5D. *FLI1* shows decreased expression via qPCR analysis in HUVECs after *FLI1* siRNA knockdown (purple) as compared to negative control siRNA knockdown (black).

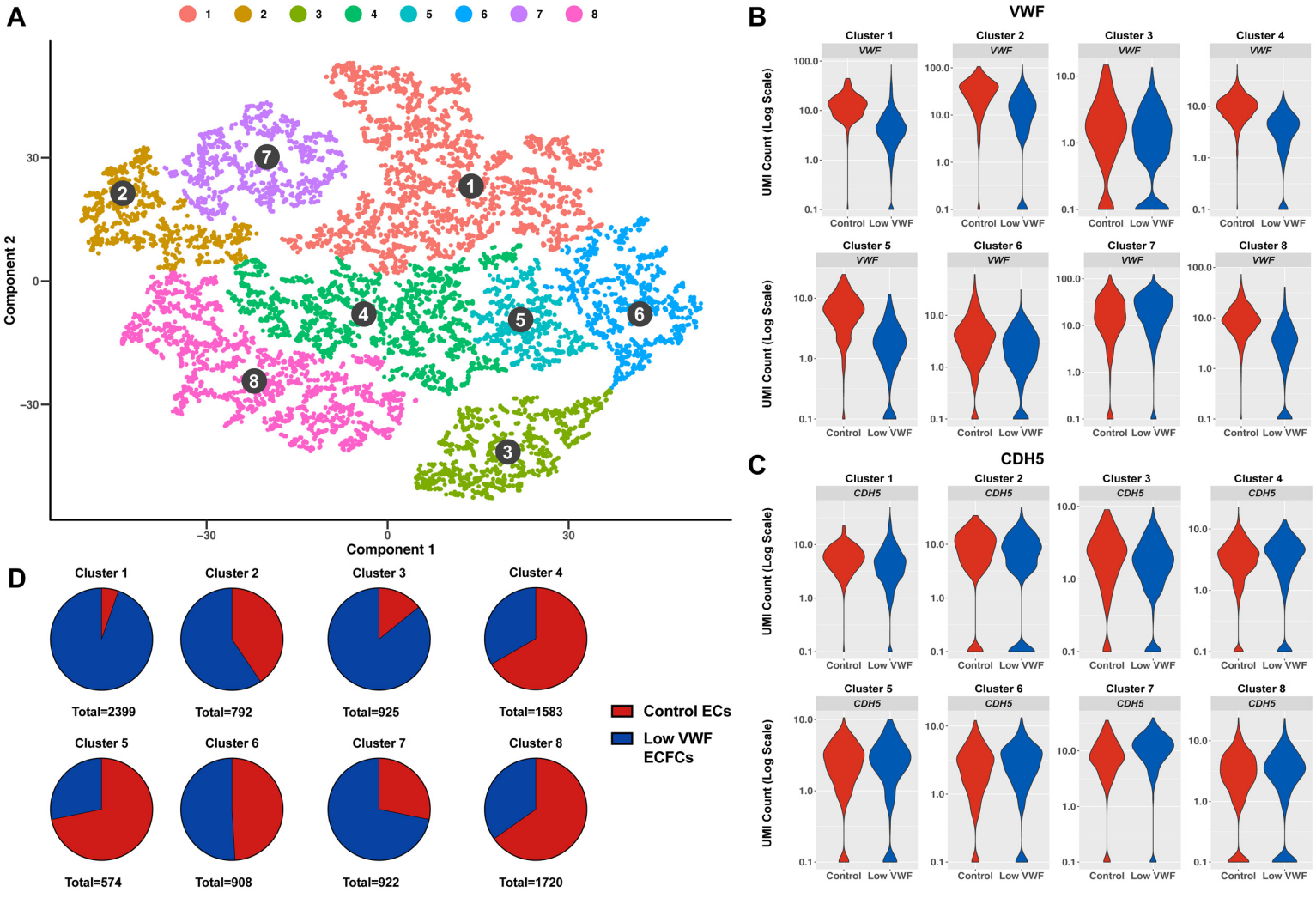
5E. *VWF* shows decreased expression via qPCR analysis in HUVECs after *FLI1* siRNA knockdown (purple) as compared to negative control siRNA knockdown (black). *VWF* siRNA knockdown is shown in pink.

5F. *VWF* promoter reporter assay expressing fluorescent eGFP in lentiviral-transduced HUVECs. After siRNA transfection with either negative control siRNA knockdown (black) or *FLI1* siRNA knockdown (purple), there is decreased reporter activity in *FLI1* siRNA as compared to negative control siRNA.

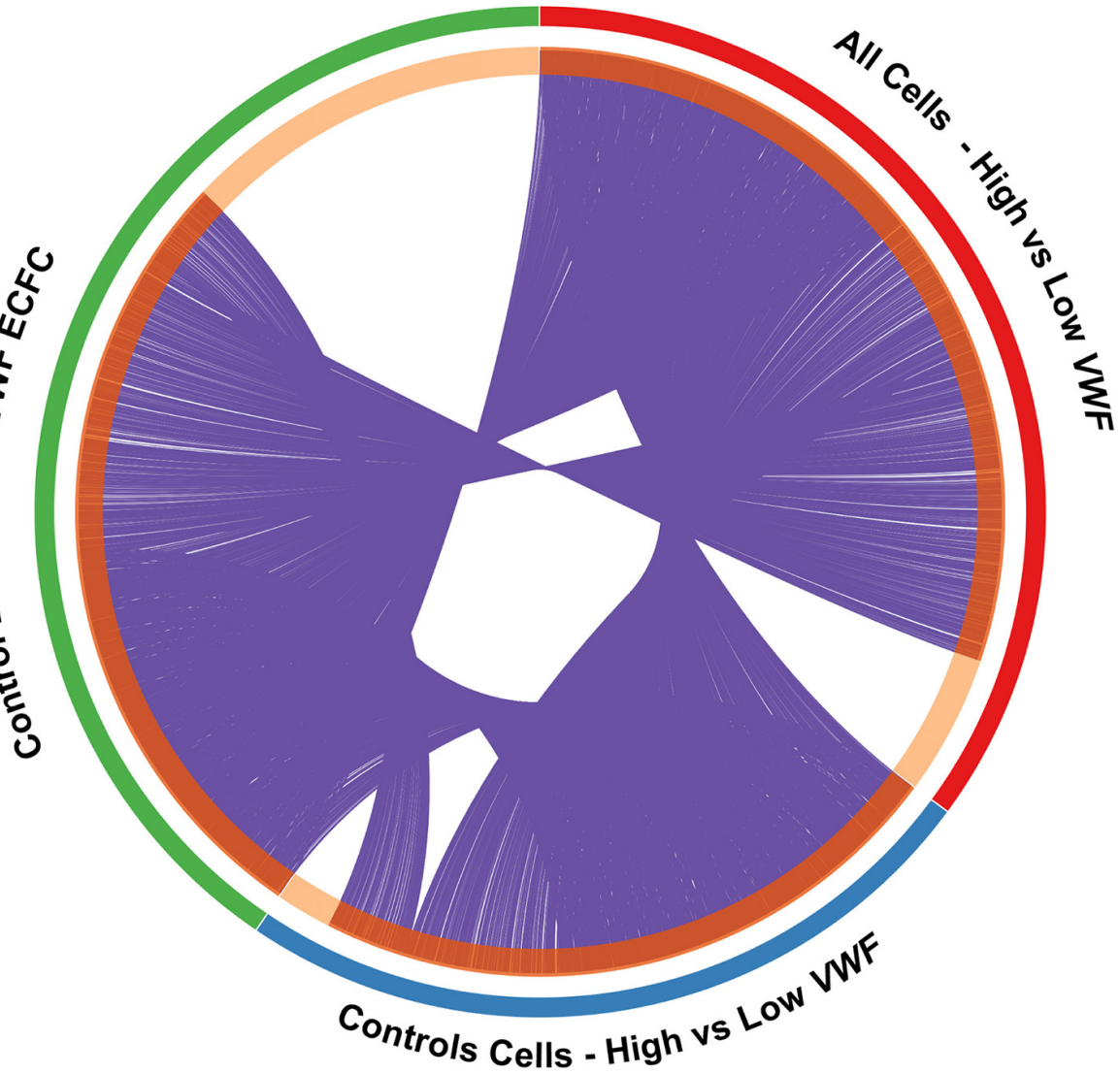
For all experiments, N>3 and p-values of significant relationships are shown as analyzed by Wilcoxon test or Kruskal-Wallis test with Dunn's multiple comparison test. Error bars represent the means ± SEM. Statistical significance is shown with asterisks as compared to control siRNA (** $p< 0.01$, *** $p<0.001$, **** $p < 0.0001$)


A
Endothelial
Non−Endothelial
Component 1
Component 2
B
Endothelial
Non-Endothelial
Low VWF ECFC28
Low VWF ECFC17
Low VWF ECFC14
Low VWF ECFC12
Low VWF ECFC11
Control ECFC24
Control ECFC3
HUVEC
Cells (Raw Number)
C
Control ECFC24
Control HUVEC
Low VWF ECFC12
Low VWF ECFC17
Control ECFC3
Low VWF ECFC11
Low VWF ECFC14
Low VWF ECFC28
D
Control ECs
Low VWF ECFCs
E
Decreased Expression in Low VWF ECFCs
Increased Expression in Low VWF ECFCs
Log10(adj. P Value)
Log2(FoldChange)
F
-log(p-value)
z-score
Protein Ubiquitination Pathway
Integrin Signaling
Mitochondrial Dysfunction
Phagosome Maturation
Oxidative Phosphorylation
Sirtuin Signaling Pathway
Actin Nucleation by ARP-WASP Complex
Remodeling of Epithelial Adherens Junctions
PI3K/AKT Signaling
Virus Entry via Endocytic Pathways
IGF-1 Signaling
Hypoxia Signaling in the Cardiovascular System
EIF2 Signaling
Unfolded protein response
Huntington's Disease Signaling
PTEN Signaling
Regulation of eIF4 and p70S6K Signaling
Ephrin Receptor Signaling
Axonal Guidance Signaling
Signaling by Rho Family GTPases
Paxillin Signaling
fMLP Signaling in Neutrophils
Ephrin B Signaling
Rac Signaling
Role of Tissue Factor in Cancer
IL-8 Signaling
Germ Cell-Sertoli Cell Junction Signaling
NRF2-mediated Oxidative Stress Response
Actin Cytoskeleton Signaling
Glucocorticoid Receptor Signaling
G
row min
row max
Low VWF ECFC12
Low VWF ECFC11
Low VWF ECFC17
Low VWFECFC14
Low VWF ECFC28
Control ECFC24
Control HUVEC
Control ECFC3
Cell_Line
id

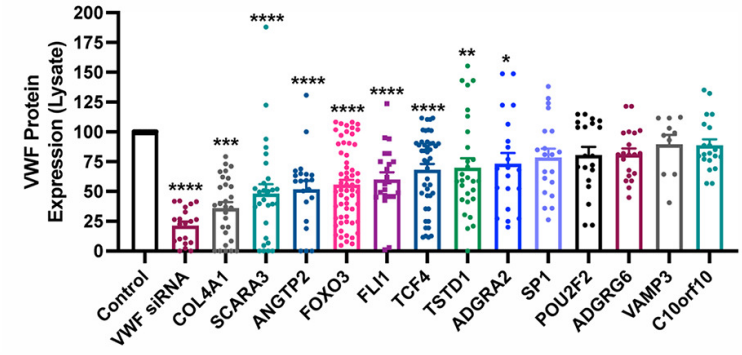

A
****
VWF Transcripts (UMI/cell)
Control ECs
Low VWF ECFCs
Control ECFC3
Control ECFC24
Control HUVEC
Low VWF ECFC11
Low VWF ECFC12
Low VWF ECFC14
Low VWF ECFC17
Low VWF ECFC28
B
High VWF Expression (39.0%)
Low VWF Expression (56.0%)
No VWF Expression (6.0%)
C
Control ECs
Normalized Density
VWF Expression (UMI/cell)
D
Low VWF ECFCs
Normalized Density
VWF Expression (UMI/cell)

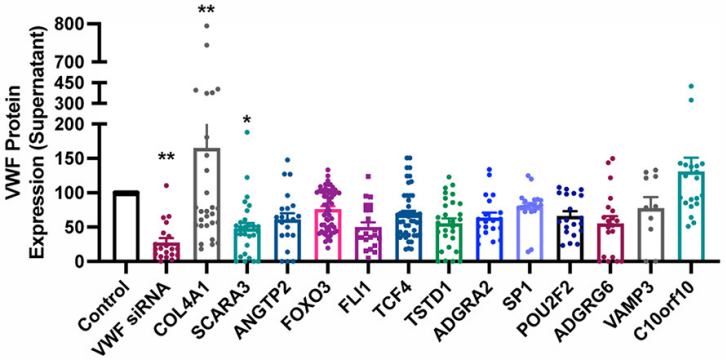

A
1
2
3
4
5
6
7
8
Component 2
Component 1
30
0
−30
B
VWF
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
UMI Count (Log Scale)
Control
Low VWF
C
CDH5
D
Total=2399
Total=792
Total=925
Total=1583
Total=574
Total=908
Total=922
Total=1720
Control ECs
Low VWF ECFCs

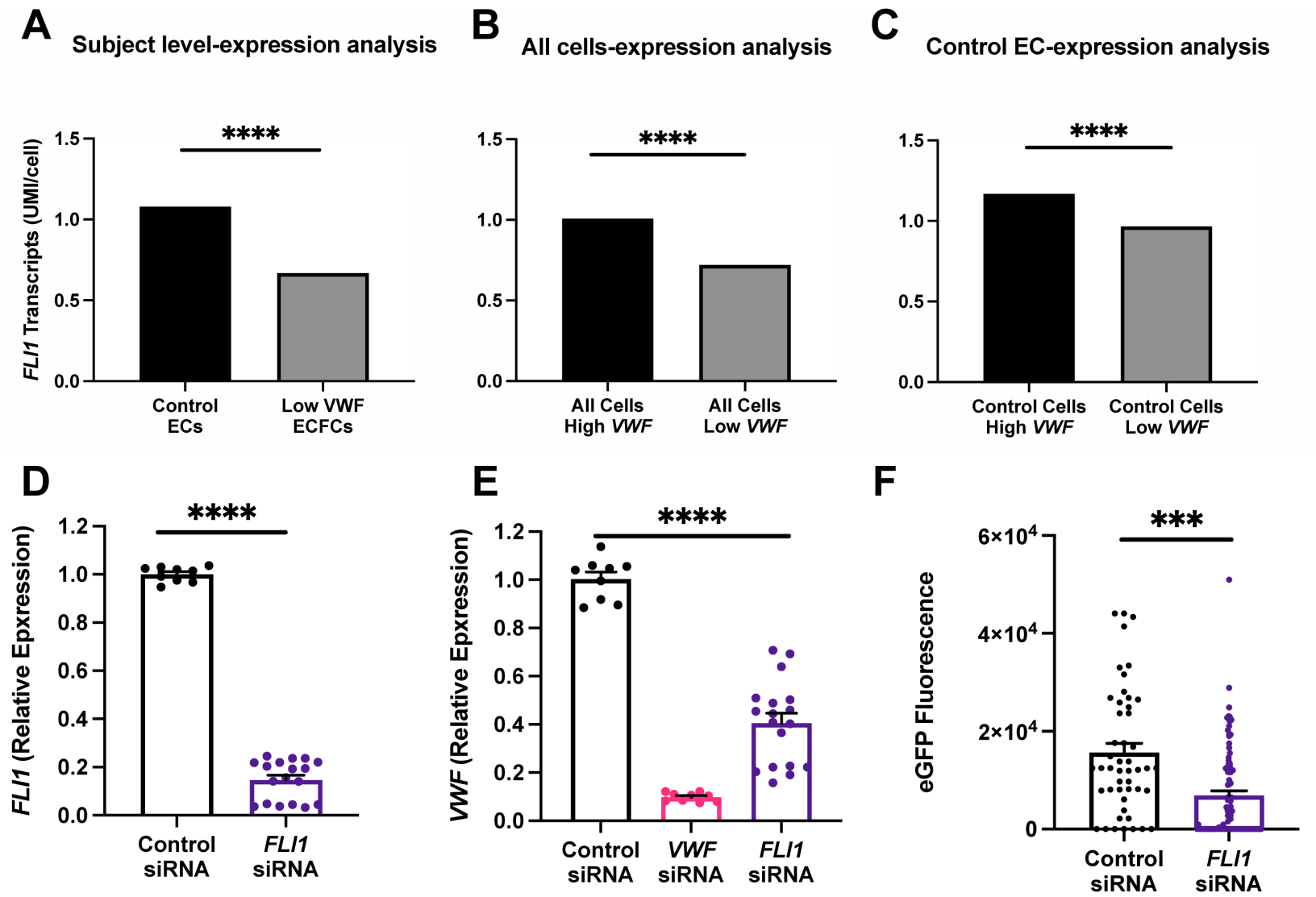
A


All Cells - High vs Low VWF
Controls Cells - High vs Low VWF
Control EC vs. Low VWF ECFC

B


VWF Protein Expression (Lysate)
200
175
150
125
100
75
50
25
0
****
***
**
*
Control
VWF siRNA
COL4A1
SCARA3
ANGTP2
FOXO3
FLI1
TCF4
TSTD1
ADGRA2
SP1
POU2F2
ADGRG6
VAMP3
C10orf10


VWF Protein Expression (Supernatant)
800
700
450
300
200
150
100
50
0
**
*
Control
VWF siRNA
COL4A1
SCARA3
ANGTP2
FOXO3
FLI1
TCF4
TSTD1
ADGRA2
SP1
POU2F2
ADGRG6
VAMP3
C10orf10


A
Subject level-expression analysis
FLI1 Transcripts (UMI/cell)
1.5
1.0
0.5
0.0
****
Control ECs
Low VWF ECFCs
B
All cells-expression analysis
****
All Cells High VWF
All Cells Low VWF
C
Control EC-expression analysis
****
Control Cells High VWF
Control Cells Low VWF
D
FLI1 (Relative Epxression)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
****
Control siRNA
FLI1 siRNA
E
VWF (Relative Epxression)
****
Control siRNA
VWF siRNA
FLI1 siRNA
F
eGFP Fluorescence
6×10^4
4×10^4
2×10^4
0
***
Control siRNA
FLI1 siRNA