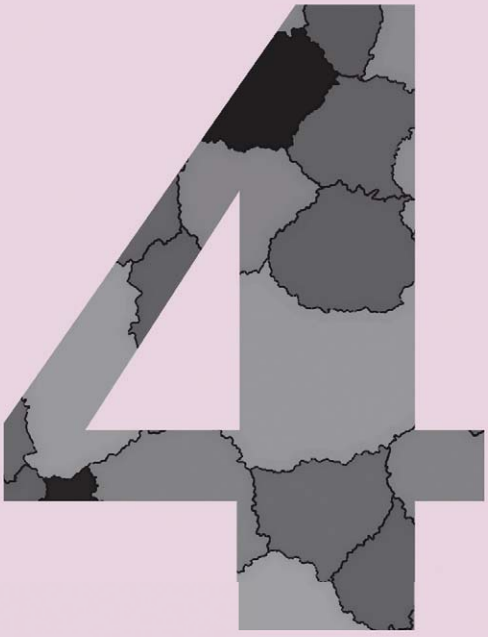
The endothelial compartment as a disease modifier in bleeding disorders

Laan, S.N.J.

Transcriptional and functional profiling identifies inflammation and endothelial-to-mesenchymal transition as potential drivers for phenotypic heterogeneity within a cohort of endothelial colony forming cells



Laan SNJ, de Boer S, Dirven RJ, van Moort I, Kuipers TB, Mei H, Bierings R, Eikenboom J

Journal of Thrombosis and Haemostasis. 2024 Jul;22(7):2027-2038.
doi: 10.1016/j.jtha.2024.03.018.

Abstract

Background

Endothelial colony forming cells (ECFCs) derived from patients can be used to investigate pathogenic mechanisms of vascular diseases like Von Willebrand disease. Considerable phenotypic heterogeneity has been observed between ECFC clones derived from healthy donors. This heterogeneity needs to be well understood in order to use ECFCs as endothelial models for disease. Therefore, we aim to determine phenotypic and gene expression differences between control ECFCs.

Methods

A total of 34 ECFC clones derived from 16 healthy controls were analyzed. The transcriptome of a selection of ECFC clones (n=15) was analyzed by bulk RNA sequencing and gene set enrichment analysis. Gene expression was measured in all ECFC clones by qPCR. Phenotypic profiling and migration speed of the ECFCs was done using confocal microscopy, followed by automated quantification of cell morphometrics and migration speed.

Results

Through hierarchical clustering of RNA expression profiles, we could distinguish two major clusters within the ECFC cohort. Major differences were associated with proliferation and migration in cluster 1, and inflammation and endothelial to mesenchymal transition in cluster 2. Phenotypic profiling showed significantly more and smaller ECFCs in cluster 1 which contained more and longer Weibel-Palade bodies (WPBs). Migration speed in cluster 1 was also significantly higher.

Conclusion

We observed a range of different RNA expression patterns between ECFC clones mostly associated with inflammation and clear differences in WPB count and structure. We developed a qPCR panel which can be used for the characterization of ECFC clones which is essential for the correct analysis of pathogenic mechanisms in vascular disorders.

Introduction

Due to their role in hemostasis, endothelial cells play a major part in many bleeding disorders, which are caused by the disruption of normal functioning hemostasis. Von Willebrand factor (VWF) is a main component of hemostasis and is produced by endothelial cells and megakaryocytes and can bind to collagen at sites of injury and mediate the formation of a platelet plug. VWF is stored in endothelial specific cigar-shaped organelles called Weibel-Palade bodies (WPB) (2, 23). These organelles can secrete their contents without stimulation to provide a steady level of VWF in the blood, thus maintaining homeostasis. WPBs secretion can also be stimulated by vascular damage to quickly increase local concentration of VWF. Von Willebrand disease (VWD) is the most common inherited bleeding disorder worldwide, occurring in roughly 1 in 100 people (24), caused by defects in concentration or structure-function of VWF (25).

One model that can be used to study pathogenic mechanisms involving the endothelium are endothelial colony forming cells (ECFCs), first described by Lin *et al.* (26). A major advantage of this model is that ECFCs can be derived from whole blood and when cultured, display endothelial characteristics such as the production of VWF, storage of VWF in WPBs, a typical cobblestone like morphology, and response to endothelial specific stimuli (27). When these cells are derived from patients, they can be used to study the pathogenic mechanisms of vascular diseases like VWD in their native environment (28-33). These studies performed patient-specific *ex vivo* analysis of endothelial cell function, endogenous production and secretion of VWF and WPB morphology.

Unfortunately, despite the advantages, there is also a challenge with using ECFCs. Substantial phenotypic heterogeneity can exist between ECFC controls from different donors, but also between clones from the same donor (5, 6). Our group has previously shown that, when comparing ECFCs categorized by their morphology, clear differences can be observed in expression of cell surface markers, proliferation, and storage and secretion of VWF (6). In the study by de Boer *et al.* (6), ECFCs were categorized into three groups. Group 1 consisted of ECFC clones with classic endothelial morphology, and group 2 and 3 consisted of larger, more elongated ECFCs. ECFCs in all groups expressed endothelial cell surface markers. However, group 1 ECFCs produced and secreted more VWF in steady state and after stimulation than groups 2 and 3. Furthermore, cell proliferation was lower in group 3. It is currently unclear what causes the heterogeneity observed among healthy control ECFCs.

In order to use ECFCs as a robust model to study pathogenic mechanisms in patients with bleeding disorders or other diseases involving endothelium and to compare with healthy donor ECFCs, it is essential to match proper control ECFCs which are characterized similarly to patient ECFCs. Use of non-characterized control ECFCs could lead to invalid conclusions when analyzing patient ECFCs. Therefore, in the current study, we analyzed differences in ECFC RNA expression, WPB count, morphology and cellular location, and migration speed of ECFCs. We found significant differences in RNA expression between ECFCs clones and these results were used to categorize the EFCFs in two distinct clusters. Between these clusters, large differences were found in WPB count, morphology and in ECFC migration speed.

Methods

ECFC isolation and culture

The study protocol for the acquisition and culturing of the ECFCs was approved by the Leiden University Medical Center ethics review board. From 16 healthy participants informed consent was obtained in accordance with the declaration of Helsinki. Participants were 18 years or older and were not diagnosed with a bleeding disorder nor known to have a bleeding phenotype. Isolation and cell culture of ECFCs was performed as described (6). In short, whole blood was obtained via venipuncture and peripheral blood mononuclear cells (PBMCs) were isolated and cultured in EGM-10 (EBM-2 Basal Medium with EGM-2 supplements & growth factors (Lonza, Basel, Switzerland or PromoCell, Heidelberg, Germany)). In general, clones appeared between days 10 and 21 and were frozen once they confluently filled 3 T75 flasks at passage 3. Multiple clones were isolated per donor, totaling 34 clones for this study. The experiments were performed on the clones at passage 5. See Supplemental Table 1 for a detailed description per clone.

RNA isolation and sequencing

All ECFC clones were cultured in 6-well plates, after they reached confluency they were kept in culture for 5-7 days. ECFC lysates were collected in 400 µl RNA lysis buffer + 4 µl β-Mercaptoethanol. RNA was isolated using RNeasy Mini Kit, according to the manufacturers' protocol (Qiagen, Hilden, Germany). 20 ng/µl of isolated RNA from 15 samples (Supplemental Table 1) was further processed by the GenomeScan facility (Leiden, The Netherlands) using the NEBNext Ultra Direction RNA library Prep Kit from Illumina (San Diego, United States). All 15 samples met the quality criteria and were selected for Bulk mRNA sequencing (polyA enriched) using Illumina NovaSeq6000. RNA sequencing (FASTQ) files were processed using the opensource BIOWDL RNA

sequencing pipeline v5.0.0 (https://zenodo.org/record/3975552) developed at the Leiden University Medical Center (Leiden, The Netherlands). This pipeline performs FASTQ preprocessing (including quality control (QC), quality trimming, and adapter clipping), RNA sequencing read alignment, read quantification, and optionally transcript assembly. FastQC (v0.11.9) was used for checking raw read QC. Adapter clipping was performed using Cutadapt (v2.10) with default settings. RNA sequencing reads' alignment was performed using STAR (v2.7.5a) on the GRCh38 human reference genome. The gene read quantification was performed using HTSeq-count (v0.12.4) with the setting "–stranded=yes". The gene annotation used for quantification was Ensembl version 105.

RNA quantification with quantitative PCR (qPCR)

RNA isolate was acquired as mentioned above. Complementary DNA (cDNA) was synthesized using SuperScript II Reverse Transcriptase (Thermo Fisher Scientific, Waltham, United States) with poly(T) primers (Sigma-Aldrich, Saint Louis, United states). Sybr Select Master Mix (Thermo Fisher Scientific) was used for qPCR which was measured on the ViiA 7 Real-Time PCR system (Thermo Fisher Scientific). GAPDH was used as housekeeping gene. Results were analyzed using the comparative Ct method. One gene panel was used on all 34 samples. The primer sequence is available as Supplemental Table 2. For analysis and creation of the heatmaps, the heatmaply package was used in R (version 4.2.1) (34). See Supplemental File 1 for an R script template that can be used to generate the heatmaps (also made available on GitHub https://github.com/Clotterdam/Laan-et-al-2023-ECFC).

Expression analysis

For the gene expression based clustering analysis, R was used. First, the edgeR package (v3.36) was used to calculate CPM (Count Per Million) of all genes in our samples. Then we selected expressed genes with a CPM higher than 1 in at least 25% of all samples (4 out of 15 samples). 12663 genes passed this filtering step. Using the dgeAnalysis package (v1.5.2), Principal Component Analysis (PCA) was performed using these 12663 genes. Plotting principal component (PC) variance showed that PC 1-4 explained most of the variance (Supplemental Figure 1A). Samples were then hierarchically clustered based on these four PCs with package cluster (v2.1.6) (Supplemental Figure 1B). The read count data of the 15 samples was labelled either as cluster 1 or cluster 2 based on the hierarchical clustering mentioned previously. Next, EdgeR (v3.36) was used to detect the differentially expressed genes between cluster 1 and cluster 2 with the trimmed mean of M values (TMM) normalization. All genes with a False Discovery Rate (FDR) adjusted p-value < 0.05 were declared significant. Gene set enrichment analysis (GSEA) was performed using the enrichplot package (version 1.18.4) and ClusterProfiler (version

4.8.2) in R on the GO:BP, GO:CC, GO:MF, KEGG and Reactome databases (35, 36). GSEA results were also visualized as a GSEA map using the enrichplot package. In the map, each node represents a significantly enriched gene-set and edge thickness represents the similarity between nodes. Node clusters were identified by the package and given a generated label. This label was later manually revised to fit the contents of the cluster.

Immunofluorescence of ECFCs and Image Acquisition

When ECFCs were plated for RNA isolation, 48-well plates (Nunclon) filled with collagen (50 µg/mL Collagen type I rat tail (BD Biosciences, Franklin lakes, United states) coated 9 mm glass coverslip (VWR, Radnor, United States) were also plated. Cells were left confluent for 5 days before fixation with 70% ethanol on ice for 10 minutes. After fixation, samples were blocked using blocking buffer (PBS; 1% bovine serum albumin (Sigma-Aldrich); 1% fetal calf serum (Gibco)). Then, samples were stained with antibodies against VWF and VE-cadherin (Supplemental Table 3 for supporting information on antibodies) diluted in blocking buffer. After, samples were stained with secondary antibodies diluted in blocking buffer and then with Hoechst in PBS, coverslips were placed on a glass slide and mounted with ProLong® Diamond Antifade Mountant (Thermo Fisher Scientific). Imaging was performed with the Imagexpress Micro Confocal System which made a tile scan (4x4) using the 63x objective without extra magnification. A z-stack was made which spanned the entire thickness of the confluent cell layer. This was transformed into a maximum Z-projection using ImageJ (version 2.3.0) (13).

Migration assay and image acquisition

Six ECFC clones were selected for the migration assay (Supplemental Table 1). We chose clones from each identified cluster 1 and cluster 2 (for both clusters n=3). These clones also belonged to each of the previously identified morphological groups 1, 2 and 3 (6) (for all groups, n=2). These clones were cultured in 48-well plates. Each clone was plated in six randomly chosen per plate. Three days post-confluency, cells were washed once with PBS and then labeled with CellTracker Green (Life Technologies) diluted 1:10,000 in 200 µL EGM-10 for 45 minutes. Three of the wells per clone were then treated with 12.5 µg/mL Mitomycin C (Sigma-Aldrich) diluted in 200 µL EGM-10 for 2 hours. The remaining wells just received EGM-10. After 2 hours, the confluent cell layer was scratched using a p100 pipet tip. Cells were washed once to remove debris. Live cell imaging was performed using the confocal AF6000 (Leica, Wetzlar, Germany) microscope with a 10x lens at 37^0C and 5% CO_2. A grid was prepared so that the same spot in the center of each well was imaged for 20 hours at 30 minutes intervals, with auto-focus correction.

Automated quantification of morphology and migration speed using CellProfiler

For the quantification of cell and organelle morphology and migration speed, CellProfiler (version 4.2.1) was used (7). We used the OrganelleProfiler (37) which is a pipeline specifically designed for the identification and quantification of cell shape, size and organelle count, shape, size and relative location in the cell. The pipeline was optimized for the antibodies and intensity in this set of tile scans. For the migration assay, a new pipeline was developed. Using the CellTracker Green signal, each cell is identified as an object. Then, cells in close vicinity to each other (confluent cells) are combined as one object. The surface area of that object was then measured. Closing speed was calculated as the increase in number of covered pixels per hour over the first 10 hours. The pipeline developed for the migration assay is available in the supplement (Supplemental File 2) and made available on GitHub (https://github.com/Clotterdam/Laan-et-al-2023-ECFC).

Statistical Analysis

Data analyses was performed using GraphPad Prism 9.3.1 (GraphPad Software, San Diego, CA, USA) if not otherwise indicated. Results with p-value < 0.05 were considered statistically significant. P-values are indicated in the figures where applicable. Unpaired T-test was performed on normally distributed data and Mann-Whitney U test was performed on not normally distributed data to compare ECFCs. A two-way ANOVA was used for the migration assay.

Results

RNA expression profile-based characterization identifies two clusters of ECFCs

A total of 34 ECFC clones derived from 16 healthy controls were analyzed, covering the previously defined phenotypic groups 1, 2 and 3 (6). Since these groups have been shown to differ in terms of surface levels of endothelial markers and storage and secretion of VWF, we also wanted to analyze differences in the RNA expression profile. Bulk RNA sequencing was performed as an unbiased method to examine transcriptional heterogeneity between healthy ECFC clones. The transcriptomes of a selection of ECFC clones (n=15) (Supplemental Table 1) were analyzed. Principal Component analysis (PCA) (Figure 1A) revealed considerable variety between the ECFC clones. Hierarchical clustering of the samples (Supplemental Figure 1) based on the PCA resulted in 2 main clusters, from here on named cluster 1 (n=11, green) and cluster 2 (n=4, blue). There is an unbalanced number of replicates between the clusters, which may cause a slight bias towards the detected differentially expressed genes to be more robust towards

cluster 1 and less robust towards cluster 2. The clusters somewhat correspond to the previously defined phenotypic groups 1 and 3 (Figure 1A). However, group 2 ECFCs do not fall specifically in either of the two clusters and showed high variation in RNA expression, likely due to this group of cells being difficult to categorize based on their morphological characteristics. Morphological group 2 ECFCs possibly represent an intermediate between cluster 1 and 2.

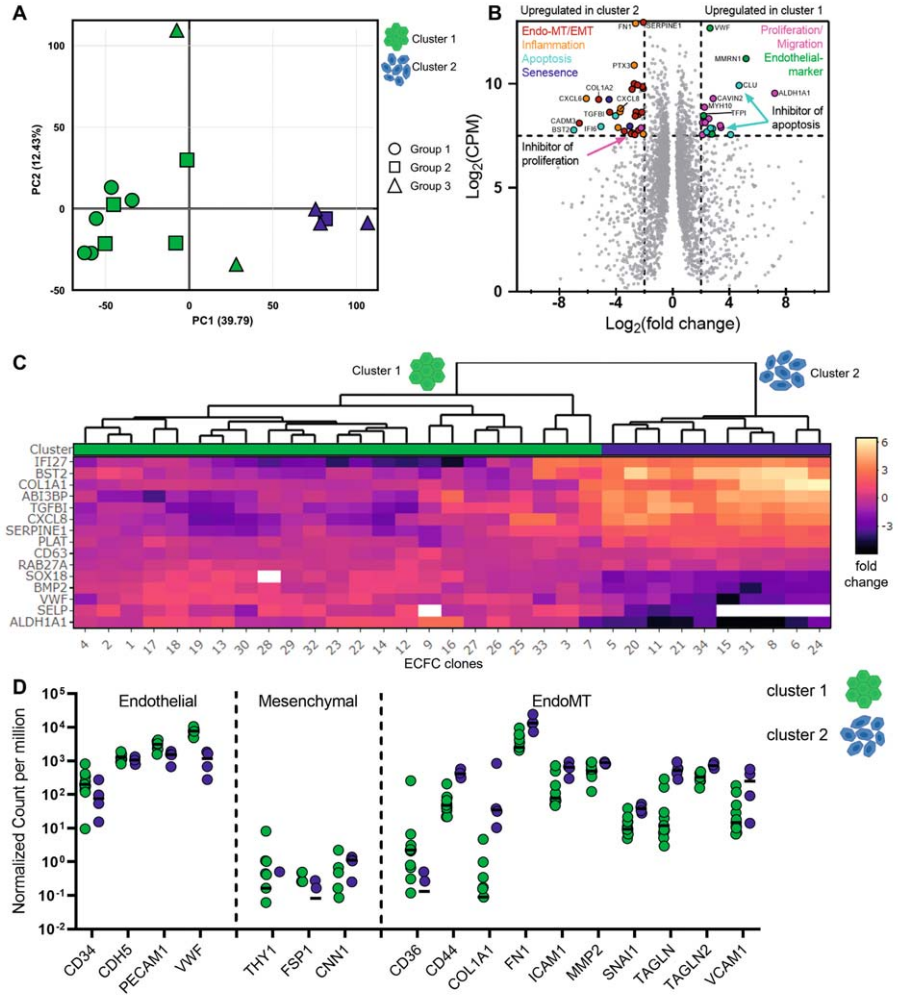


Figure 1. RNA expression analysis showing differential gene expression between ECFC clones. A) PCA plot of 15 ECFC clones, color-coded by their associated cluster as measured by hierarchical clustering. Circles, squares and triangles indicate the original categorization of ECFCs based on cell morphology in group 1, group 2 and group 3 respectively (13). B) Bulk RNA sequencing transcriptome analysis shows differential gene expression between ECFC clusters. The volcano plot shows significantly upregulated genes (grey, adjusted P-value < 0.05) between

cluster 1 and cluster 2 (3758 genes). Dashed lines represent the threshold where fold change was > 2 and counts per million (CPM) were > 7.5. Genes involved with proliferation and migration are shown in pink and endothelial markers are indicated in green. Genes involved in endothelial to mesenchymal transition (EndoMT) and epithelial to mesenchymal transition (EMT), inflammation, apoptosis, and senescence are shown in red, orange, cyan and blue respectively. Gene names of genes with the highest fold change and count per million are displayed. C) Heatmap showing the log nFOLD difference of the RNA expression with the median of 33 ECFC clones as measured by qPCR. Hierarchical clustering of the ECFCs is shown by the dendrogram (top). All ECFCs were tested on a panel of genes (Supplemental Table 2). White squares marks data points that could not be measured. D) Normalized RNA count per million (CPM) for endothelial, mesenchymal and EndoMT markers. The y-axis is shown as a log(10) scale. Based on differential gene expression analysis between clusters, PECAM1, VWF and all EndoMT markers were significant (adjusted p-value). CD34, CDH5 and all mesenchymal were not significant. Each dot represents the CPM per clone.

Gene expression differences between cluster 1 and cluster 2 ECFCs were analyzed. From all samples, 11,817 genes were measured and 3,758 genes showed significant differential expression (Figure 1B). Genes with a fold change higher than 2 and a count per million (CPM) higher than 7.5 were further investigated. This cut-off resulted in 54 genes of interest (Supplemental Table 4). Fold change is reported as the difference in expression in cluster 1 compared to cluster 2. Genes with a positive fold change are thus upregulated in cluster 1. Genes of interest involved in processes like endothelial to mesenchymal transition (EndoMT), inflammation, apoptosis, and senescence were upregulated in cluster 2 ECFCs. In contrast, upregulated genes in cluster 1 ECFCs are associated with proliferation and migration. Cluster 2 ECFCs also showed significant upregulation of Collagen Type I Alpha 1 Chain and C-X-C Motif Chemokine Ligand 8, and cluster 1 ECFCs showed upregulation of SRY-Box 18, Ephrin B2 and Thrombomodulin which corresponds with previous findings of de Jong *et al.* (5).

We designed a minimal qPCR panel with the aim to easily characterize the clones based on their gene expression. The panel was created based on the results of the bulk RNA sequencing and consists of genes which showed the strongest fold change, highest CPM and biological relevance to characterize ECFC clones (Supplemental Table 2). The qPCR panel was used to analyze the RNA expression of all clones. Per clone, gene fold change was compared to the median of all ECFC clones, as shown in a heat map (Figure 1C). Considerable variation between the clones can be observed, especially for the genes Interleukin-8 (*CXCL8*), Interferon alpha-inducible protein 27 (*IFI27*), Bone Marrow Stromal Cell Antigen 2 (*BST2*) and Collagen type I alpha (*1COL1A1*) which where downregulated in cluster 1 ECFCs. Whereas VWF, P-selectin (*SELP*) and aldehyde dehydrogenase 1 family member A1 (*ALDH1A1*) were upregulated in cluster 1. The hierarchical clustering of the clones based on this selection of genes results in the

same clusters as those observed in the PCA plot. This indicates that the qPCR panel can accurately categorize ECFC clones using a minimal list of targets.

Differential expression indicates a role for EndoMT. To further substantiate this claim, we have highlighted endothelial markers (*VWF, CD34, CD144* and *PECAM1*) (38), mesenchymal markers (*CNN1, THY1* and *FSP1*) (39), and early and late EndoMT markers (*TAGLN, CD44, FN1, 1COL1A1, MMP2, VCAM, ICAM1* and *SNAI1*) (39, 40) from the RNA sequencing results (Figure 1D). All the ECFCs show strong expression of endothelial markers although cluster 2 cells show significantly decreased expression of endothelial markers *PECAM1* and *VWF*. It was also observed that the mesenchymal markers *CNN1, THY1* and *FSP1* had very low expression in almost all samples which were not significantly different between clusters, indicating that the cells are not fully mesenchymal. Interestingly, the EndoMT markers where expressed by all ECFCs and were all significantly upregulated in cluster 2 compared to cluster 1. This further emphasizes that ECFCs are still endothelial and not fully mesenchymal cells and that the EndoMT pathway plays a significant role in the heterogeneity between clones. These findings give an indication of the differences between clones on an expression level, but additional experimental evidence is needed to confirm this.

Inflammation and endothelial to mesenchymal transition pathways are differentially regulated in cluster 1 and cluster 2 ECFCs

To formally identify and prioritize relevant gene sets associated with the observed differences between the cluster 1 and cluster 2 ECFCs (Figure 2), we employed Gene Set Enrichment Analysis (GSEA). GSEA is a powerful computational approach that offers several advantages in the analysis of bulk RNA sequencing data like reducing the impact of random noise in large-scale transcriptomic datasets (41). We applied GSEA to the Gene Ontology (GO) Biological Process (BP), Cellular Component (CC) and Molecular Function (MF) databases. GSEA from the GO:BP (Figure 2A) demonstrated significant differences (q-value < 0.05) with a positive normalized enrichment score (NES) (meaning an upregulation in cluster 1 ECFCs) in the gene set "regulation of endothelial cell migration" (NES = 1.56), and a negative NES (meaning a downregulation in cluster 1 ECFCs) in "inflammatory response" (-1.60), "extracellular matrix organization" (-1.69), and "endodermal cell differentiation" (-2.20). Furthermore the "epithelial to mesenchymal transition" (-1.57) gene set showed borderline significant differences (q-value of 0.06). This highlights the potential variation in cytokine regulation, migration and differentiation.

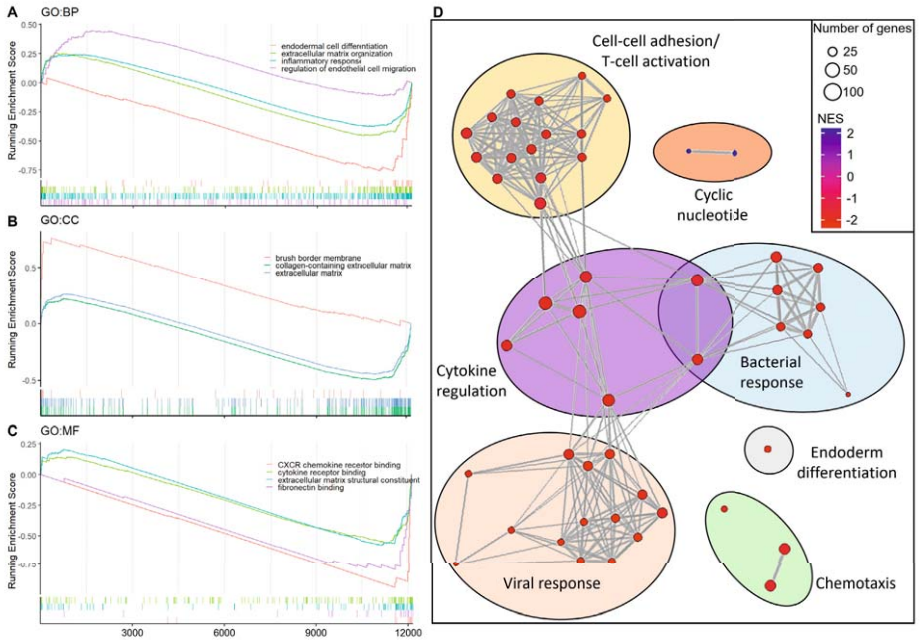


Figure 2. Gene Set Enrichment Analysis of cluster 1 versus cluster 2 ECFC RNA expression. (A-C) GSEA showed the significantly enriched gene sets (FDR < 0.05) between cluster 1 and cluster 2 ECFCs for various databases; GO:BP (A), GO:CC (B) and GO:MF (C). D) GSEA map of the top 50 gene sets from the GO:BP database constructed with the enrichplot package. Red nodes indicate gene sets with a negative NES while blue indicates a positive NES score. Thickness of lines between nodes correspond with the similarity between gene sets. Abbreviations: Normalized enrichment score (NES); Genome Ontology (GO); Biological Processes (BP); Cellular Component (CC); Molecular Function (MF).

Additionally, GSEA from the GO:CC (Figure 2B) and GO:MF (Figure 2C) databases resulted in gene sets scoring negatively for "collagen-containing extracellular matrix" (-1.90), "extracellular matrix" (-1.78), "CXCR chemokine receptor binding" (-2.12), "cytokine receptor binding" (-2.16), "extracellular matrix structural constituent" (-1.96) and "fibronectin binding" (-2.01). The Gene set "Brush border membrane" had a positive NES (1.94). Genes often participate in multiple pathways and GSEA can thus yield large numbers of broadly overlapping gene sets. Therefore we collapsed redundant pathways into single functional or biological themes and created a GSEA map for the top 50 enriched GO:BP gene sets (Figure 2D), thereby further aiding in interpretation. Clusters of gene sets indicate that most negatively enriched gene sets are associated with inflammation and immune responses, either to viruses or bacteria. These findings collectively provide a view of the transcriptional distinctions between the clusters,

offering novel insights into the underlying molecular mechanisms regulating their distinct characteristics.

Quantitative differences in cell and WPB morphology between ECFC clusters

RNA expression showed a significant difference in genes associated with cell proliferation and VWF production. It has been shown that VWF production is directly linked to the length of WPBs (14). Previous research has also shown significant differences in VWF protein production, cell size and proliferation between ECFC clones (5, 6). Therefore, we imaged cluster 1 and cluster 2 ECFC clones (Figure 3A) to quantify the cell count, cell size and shape number (Figure 3B-D). To analyze WPB count and morphology we also quantified their eccentricity and maximum ferret diameter as an approximation of roundness and length respectively (Figure 3E-H), using a specialized CellProfiler pipeline for automated WPB identification and quantification (37). Data is shown as mean ± SD. Cell count was significantly lower in cluster 2 ECFCs (n = 152.90 ± 109.60) than in cluster 1 (n = 409.90 ± 164.80, p=0.0003). As endothelial cells form a confluent layer, it follows that we observed a larger cell size in cluster 2 ECFCs (4490 μm^2 ± 2799) compared to cluster 1 ECFCs (1529 μm^2 ± 684.90, p=0.0006). We also observed that cluster 2 ECFCs were significantly rounder (0.77 ± 0.02) than cluster 1 ECFCs (0.83 ± 0.04, p=0.0006) as measured by their eccentricity. Furthermore, WPB count per cell was significantly lower in cluster 2 ECFCs (n = 25.40 ± 30.40 vs. 118.00 ± 44.40, p=0.0001). WPB eccentricity and length of the WPBs were measured and WPBs were both rounder and shorter in cluster 2 ECFCs (0.66 ± 0.03 vs. 0.72 ± 0.05, p=0.0024 and 0.76 µm ± 0.12 vs. 0.97 µm ± 0.26, p=0.017 respectively). Finally, the relative distance of the WPBs to the nucleus in percentage was measured. We found that WBPs of cluster 1 ECFCs (56.48% ± 7.20) tend to locate more to the periphery of the cell than cluster 2 ECFCs which seem to locate relatively closer the nucleus (46.09% ± 13.79, p=0.06). This is likely explained by the larger size of the cluster 2 ECFCs. Collectively, these findings could explain the reduced VWF production and secretion observed in previous research (6).

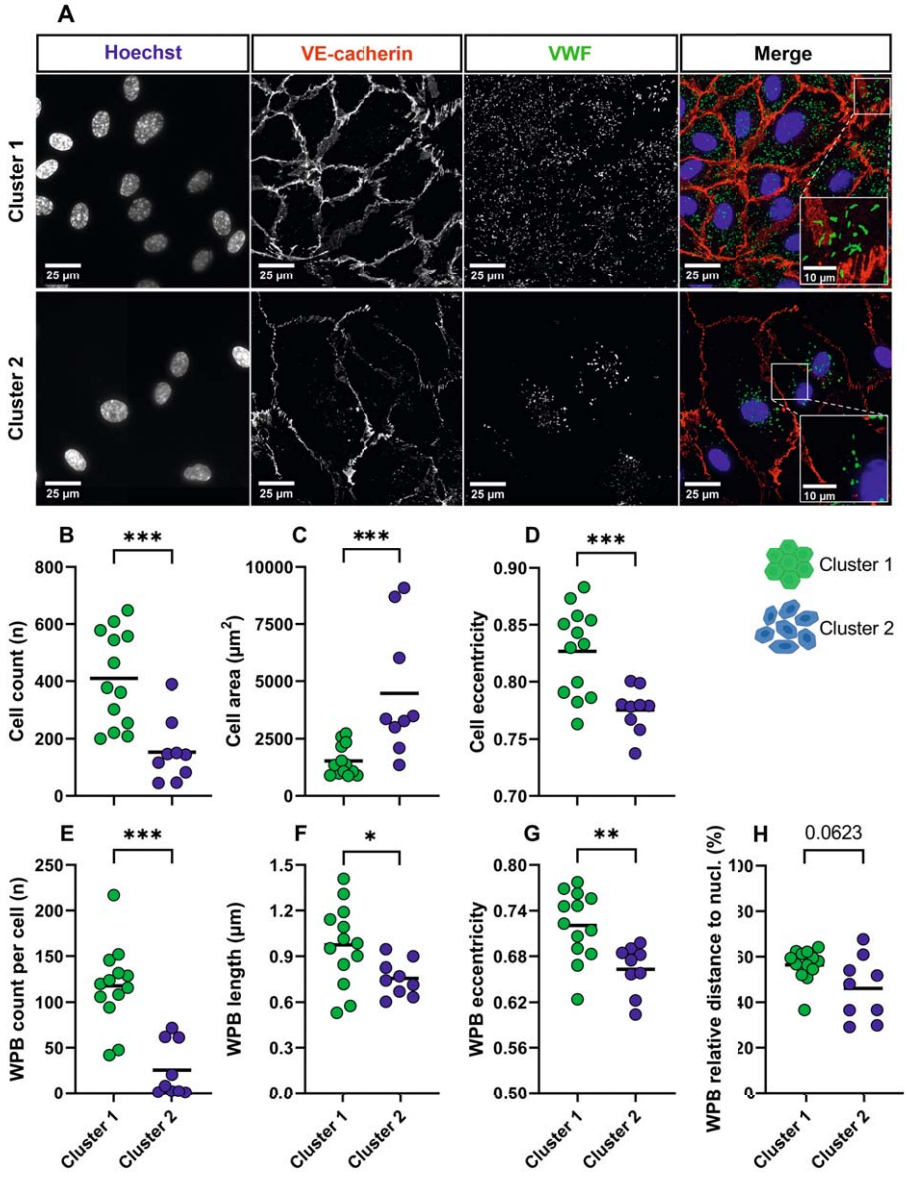


Figure 3. Morphological differences of cell and organelle count and shape. Phenotypic profiling of the ECFCs was done using tile scans (707381 µm2). ECFC clones were divided by hierarchical clustering based on RNA expression in cluster 1 (n=13) and 2 (n=9). A) Representative confocal images of ECFCs stained with Hoechst (blue) and antibodies against VE-cadherin (red) and VWF (green) from cluster 1 (top) and cluster 2 (bottom, scale bar represents 25 µm). The white box shows a 2.5x zoom in of the merge (scale bar represents 10 µm). Images were taken with a 63x times objective. B) Cell count per surface area of the tile scan. C) Median surface area of cells per clone. D) Median cell eccentricity. E) WPB count per cell per ECFC clone. F) Median WPB length per ECFC clone. G) Median eccentricity of WPBs per ECFC clone. H) Distance of the

WPB to the nucleus relative to their position in the cell in percentage. Values are median per clone. Line shows median. Unpaired T-test was performed on normally distributed data (B, D, F, G and H). Mann-Whitney U test was performed on not normally distributed data (C and E); * p <0.05, ** p < 0.01, *** p <0.001.

Decreased cell migration in cluster 2 ECFCs

According to the GSEA, genes associated with regulation of endothelial cell migration were differentially expressed between cluster 1 and cluster 2. To determine the rate of migration of the ECFCs, a scratch assay was performed three times on a selection of cluster 1 and cluster 2 ECFCs (n=3 per cluster) and closing speed was quantified (Figure 4A). The mean cell count per cluster at T0 was 418.01 ± 52.25 in cluster 1 and 252.06 ± 82.43 in cluster 2. We observed that the closing speed in the first 20 hours after the scratch was significantly higher in cluster 1 ECFCs (29143.67 ± 6713.37) than in cluster 2 ECFCs (13889.50 ± 1278.76, p=0.0097) (Figure 4B). Furthermore, the same effect was seen when the potential contribution of proliferation to the closing speed was excluded by inhibition of proliferation using Mitomycin C (26901.67 ± 6882.52 vs 12610.83 ± 399.18, p=0.0138). This indicates that the difference in closing speed is caused mostly by the migration capabilities of the cells within this time frame. This is in line with previous research where proliferation between clones of different morphological groups was analyzed (6). There, the phenotypic groups of ECFCs showed no significant difference in proliferation in the first 24 and 48 hour period, but group 1 did show increased proliferation after 48 hours in culture.

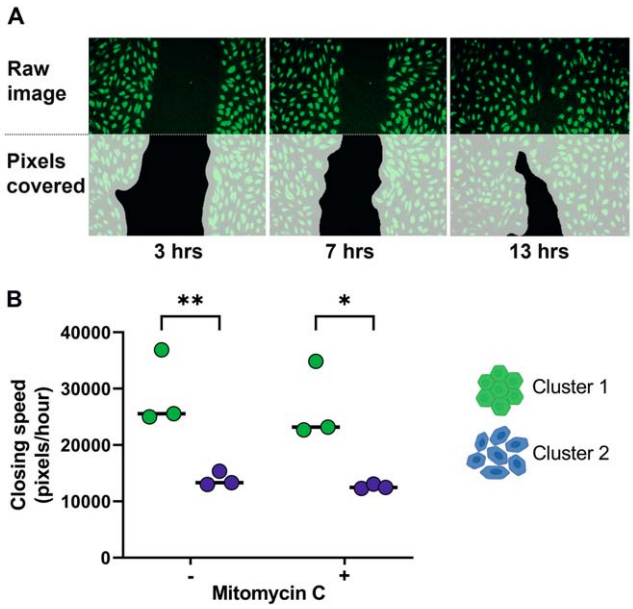


Figure 4. Delayed migration of cluster 2 ECFCs. ECFCs were stained with a live cell marker (Cell tracker), scratched and then imaged every 30 minutes. A) Top, representative image of the scratch at three time points in a cluster 1 ECFC clone. Bottom, a graphic representation of the calculated pixel coverage in white as calculated by a CellProfiler pipeline. Average pixel closing speed per hour was quantified. B) Mean closing speed of three cluster 1 ECFCs and three cluster 2 ECFCs (showing the average of three measurements). On the left without (-) and on the right with (+) inhibition of proliferation by Mitomycin C. Statistical analysis by two-way ANOVA, * $p < 0.05$, ** $p < 0.01$.

Discussion

ECFCs have been used to study the pathogenic mechanisms of various diseases *in vitro* (29, 32, 33, 42). However, considerable phenotypic heterogeneity has been observed in ECFC clones isolated from healthy controls (5, 6), which may preclude unambiguous interpretation of such work. Therefore, we aimed to determine phenotypic and gene expression differences between control ECFCs. In this study, we found that 2 major clusters of ECFCs could be discerned that, in all our subsequent morphological and functional analyses, were phenotypically distinct. Cluster 1 ECFCs are smaller cells that contain large numbers of elongated WPBs and show high migration capacity, whereas cluster 2 ECFCs are large, contain fewer WPBs that are also significantly shorter and display a reduced migratory potential. Elongated morphology of WPBs correlates with their secretion competence as well as the hemostatic potential of their main cargo protein VWF (3), while abnormalities in WPB size and shape can be a direct consequence of the pathogenic mechanisms that underpin

some forms of VWD (32, 33, 43). Moreover, it was reported that ECFCs isolated from VWD patients also have alterations in their migratory and angiogenic potential (29, 31, 44). Our observation that these key parameters already show significant differences between the 2 phenotypic clusters within healthy controls highlights the need for care when interpreting data obtained with ECFCs from healthy as well as diseased subjects. It also offers a potential strategy to minimize the impact of ECFC heterogeneity on experimental results (see below).

Despite unbalanced numbers of replicates between clusters, the high number of detected differentially expressed genes was sufficient to detect biological process changes. Cluster 1 ECFCs showed upregulation of genes associated with proliferation and migration, while cluster 2 ECFCs have upregulated genes associated with inflammation, senescence and apoptosis. This included *Pannexin 1* (data in repository), a senescence marker that was recently found to modulate angiogenic activities and cellular activity in ECFCs (45). RNA expression profiling also showed that *TGFBi, TGFB2, BMP2* and *SMAD1* were upregulated in cluster 2 which are EndoMT/EMT associated genes (39). EndoMT transforms endothelial cells into mesenchymal cells, leading to reduced expression of endothelial markers, increased extracellular matrix proteins, and loss of endothelial functions (46-48), including reduced synthesis of VWF (49). Our transcriptional, morphological and functional data fit with the notion that cluster 2 ECFCs are in the process of EndoMT or are transitioning towards mesenchymal cells. Transcriptional analysis of ECFCs by Kutikhin *et al.* showed that ECFCs overexpress endothelial markers *NRP2, NOTCH4* and *LYVE1* when compared to human coronary artery endothelial cells and human umbilical vein endothelial cells (50). These markers were also strongly expressed in our data. Interestingly, *LYVE1* was expressed in all samples, but was strongly upregulated in cluster 1 ECFCs, suggesting cluster 1 ECFCs have more potential to differentiate into the lymphatic endothelial lineage. Higher levels of *TGFB2* in cluster 2 which has been shown to induce EndoMT and downregulate expression of *LYVE1* may also explain the difference in expression (51).

Why some ECFC isolations result in clones that correspond to cluster 2 clones with EndoMT characteristics, while also yielding cluster 1 clones that have not progressed to that state is currently unclear. It has been shown that EndoMT can be regulated by epigenetic mechanisms (52), which could have led to some of the circulating cells from which ECFC clones originate having been primed upfront towards generating cluster 2 ECFCs. The origin of circulating ECFCs remains unclear. Tura et al. showed that ECFCs are likely not originating from bone marrow but are derived from a CD34(+), CD133(-), CD146(+) cell fraction potentially arising from vessels (53). The transcriptome of ECFCs closely resembles that of cultured microvascular endothelial cells and HUVECs (54). Heterogeneity between clones may arise from different vascular beds, influenced by specific microenvironmental cues (55). Lin et al. identified a fraction of $CD34^{bright}$ and $PROCR^{+}$ umbilical vein cells as a potential ECFC origin.

Interestingly, the genes used by Lin et al to characterize this subset (PROM1, PTPRC, CDH5, PECAM1, MCAM and FLI1) are also expressed similarly by the ECFCs used in our study (56) (see repository). This makes it likely that they are derived from the same pool. Lack of differential expression of these genes between ECFC clusters suggests differentiation occurred after isolation.

Multiple studies have shown that inflammation can cause EndoMT in primary endothelial cells (57-59). This is in line with our findings as the transcriptional profiling and the subsequent GSEA showed a strong upregulation of inflammation and immune response pathways in cluster 2 ECFCs. Collectively, these data suggest that cluster 2 ECFCs create, or are a result of, an inflammatory environment. Additionally, our data shows that expression of the pro-inflammatory cytokine interleukin 8 (*IL-8*) is significantly upregulated in cluster 2 ECFCs. Medina *et al.* showed that endothelial cells after *ex vivo* expansion became enlarged and senescent in an *IL-8* dependent manner (60). We speculate that this represents an autocrine/paracrine inflammatory loop that can initiate or perpetuate the transdifferentiation into cluster 2 ECFCs. Whether this is an ongoing process in which even cluster 1 ECFCs are destined to eventually become cluster 2 is not clear from our data, since the bulk RNA sequencing analysis that we performed is unable to identify single cell differences within clones. Longitudinal studies and single cell RNA sequencing analysis will be necessary to reveal whether ECFC clones are homogenous populations or in various stages of transition.

Lastly, the phenotypic differences in ECFCs could be attributed to isolation and culture conditions, influenced by factors such as prolonged culture times, frequent media changes, and clonal expansion from a single cell. Previous studies have shown the impact of variables on the characteristics of ECFC clones, such as day of first appearance of ECFCs (6), passaging (60), time in culture (61, 62) and as a result, replicative stress that some ECFCs may have experienced during expansion. A number of questions remain regarding the unpredictability of ECFC isolations. Firstly, it has been observed that some isolations from donors or certain disease phenotypes yield no clones, so termed “zero colonies” (63). For example, ECFC isolation from patients with VWD type 3 is possible but with low success rate (29), suggesting a role for VWF in this process. Secondly, donor age also seems to influence outcome as isolations from children (0 to 10 years) yield significantly more ECFCs than adult isolation, but no differences are seen between adults ranging from 20-73 years of age (63). The donors included in the current study were all adults and no age-related differences were observed in this study. Ongoing efforts by the ISTH SSC Vascular Biology aim to standardize ECFC isolation and culturing, proposing recommendations for seeding density, passaging, and clone expansion to reduce variation between laboratories and clone heterogeneity (63, 64). Whether these recommendations will favor the emergence of relatively more cluster 1 or 2 ECFCs remains to be seen.

This study provides crucial insight into the heterogeneity of ECFC clones derived from healthy donors. Whether this heterogeneity between ECFC clones relates to a predisposition within the cell of origin, or if it is introduced as a result of the inherent variations associated with primary cell isolation and culture conditions, remains unclear. In a previous, multicenter study using ECFCs isolated from Dutch and Canadian donors in 2 separate laboratories we showed that, using classification of ECFC clones based on morphology, we could distinguish 3 separate groups that also differed in terms of proliferation, VWF secretion and expression of cell surface receptors (6). We now show that an even further and simpler dimensionality reduction can be achieved using qPCR profiling. While this has only been performed in our laboratory, it broadly aligns with the previous classification of the Dutch and Canadian ECFCs (6). This highlights that these discrete phenotypic clusters are probably not restricted to ECFC isolations from our laboratory, but may also be present within ECFC collections from other investigators. Whether our strategy to dichotomize ECFCs into two distinct clusters can be more generally applied as a solution for phenotypic heterogeneity that is observed by other investigators, will need to be validated in other labs. This should include standardized methods such as proposed by the SSC (65), to rule out the effect of experimental variation during isolation and ex vivo maintenance of ECFCs. Furthermore, it is vital to acknowledge that there are other key aspects of ECFCs not addressed in this study like angiogenic capacity, proliferation, apoptosis or endothelial barrier function. For studies that aim to use ECFCs with such phenotypic readouts it is thus important to confirm that the phenotypic variability that has been observed for those aspects (44) also relates to the 2 ECFC clusters that we identify.

In order to use the ECFCs as a robust model to study pathogenic mechanisms, it is essential that one takes the phenotypic heterogeneity into account in the experimental design. This ensures that findings are not incorrectly attributed to pathogenic mechanisms rather than phenotypic heterogeneity in ECFCs. So far, no objective criteria were available that could be used to stratify ECFCs for this purpose. The benefit of this study is that we present a minimal qPCR panel that can be used as a tool to pre-characterize and dichotomize clones during the isolation workflow into two ECFC subsets, each with distinct morphological and migratory features. The relatively small number of genes that need to be screened, combined with the wide availability of qPCR platforms means that this should be a quick and cost-effective tool that is accessible for all laboratories that study ECFCs. Furthermore, qPCR offers an unbiased approach to pre-select clones, without selection based on outcome parameters. The raw qPCR data of the panel of 34 ECFC clones is made readily available and can be used to aid characterization of small numbers of samples (data in repository). When applying the qPCR panel to match ECFC clones, we recommend to select neighboring clones after hierarchical clustering, or to select clones which, in gene expression, don't differ more than 1 fold change from one another. Classification of ECFC clones provides a rationale to select

matching ECFCs in experimental comparisons. Owing to, among others, their favorable growth characteristics, one would preferentially compare cases with controls using cluster 1 ECFCs. In cases where only cluster 2 ECFCs are available, for instance with rare patients with persistent low yield of ECFCs, this classification can help to minimize the effect of phenotypic variability on experimental outcome by selecting matching cluster 2 control ECFCs. Finally, this knowledge offers an excellent platform for follow-up research to be performed. Our data suggests a strong role of an inflammatory mechanism that could cause, or be the result of, the differences between ECFCs clones. Further understanding of the cause of this inflammatory milieu could lead to improved standardization of the isolation and culturing protocol.

Acknowledgements

The SYMPHONY consortium, which aims to orchestrate personalized treatment in patients with bleeding disorders, is a unique collaboration between patients, health care professionals, and translational and fundamental researchers specializing in inherited bleeding disorders, as well as experts from multiple disciplines (22). It aims to identify the best treatment choice for each individual based on bleeding phenotype. To achieve this goal, work packages (WP) have been organized according to 3 themes (e.g. Diagnostics [WPs 3 and 4], Treatment [WPs 5-9], and Fundamental Research [WPs 10-12]). Principal investigator: M.H. Cnossen; project manager: S.H. Reitsma.

Beneficiaries of the SYMPHONY consortium: Erasmus University Medical Center-Sophia Children's Hospital, project leadership and coordination; Sanquin Diagnostics; Sanquin Research; Amsterdam University Medical Centers; University Medical Center Groningen; University Medical Center Utrecht; Leiden University Medical Center; Radboud University Medical Center; Netherlands Society of Hemophilia Patients (NVHP); Netherlands Society for Thrombosis and Hemostasis (NVTH); Bayer B.V., CSL Behring B.V., Swedish Orphan Biovitrum (Belgium) BVBA/SPRL.

Funding by SYMPHONY: NWO-NWA.1160.18.038 (received by SNJL, IvM, RB and JE), Landsteiner Foundation for Blood Transfusion Research, Grant Number:1707. (received by RB) and Landsteiner Foundation for Blood Transfusion Research, Grant/Award Number: 1852 (received by SB)

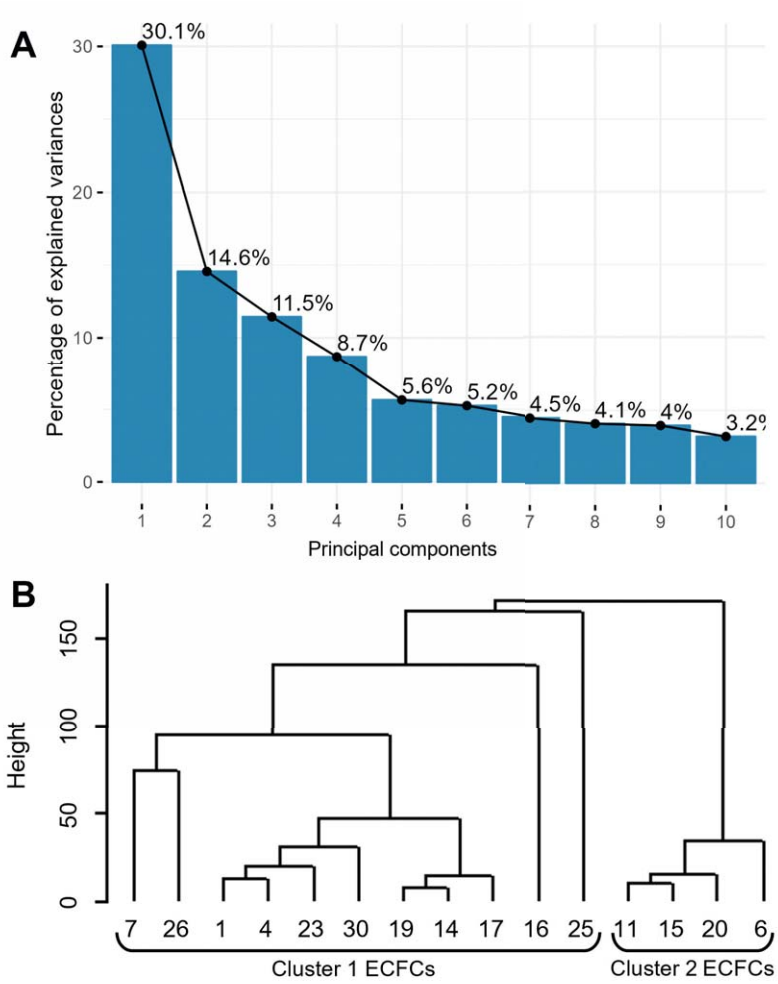
All data files are available in a Zenodo repository (https://zenodo.org/records/10422769).

Reference list

1. Valentijn KM, Sadler JE, Valentijn JA, Voorberg J, Eikenboom J. Functional architecture of Weibel-Palade bodies. Blood. 2011;117(19):5033-43.
2. Hordijk S, Carter T, Bierings R. A new look at an old body: molecular determinants of Weibel-Palade body composition and VWF exocytosis. J Thromb Haemost. 2024.
3. Murray EW, Lillicrap D. von Willebrand disease: pathogenesis, classification, and management. Transfus Med Rev. 1996;10(2):93-110.
4. Leebeek FWG, Eikenboom JCJ. Von Willebrand's Disease. New England Journal of Medicine. 2016;375(21):2067-80.
5. Lin Y, Weisdorf DJ, Solovey A, Hebbel RP. Origins of circulating endothelial cells and endothelial outgrowth from blood. J Clin Invest. 2000;105(1):71-7.
6. van den Biggelaar M, Bouwens EA, Kootstra NA, Hebbel RP, Voorberg J, Mertens K. Storage and regulated secretion of factor VIII in blood outgrowth endothelial cells. Haematologica. 2009;94(5):670-8.
7. Berber E, James PD, Hough C, Lillicrap D. An assessment of the pathogenic significance of the R924Q von Willebrand factor substitution. J Thromb Haemost. 2009;7(10):1672-9.
8. Groeneveld DJ, van Bekkum T, Dirven RJ, Wang JW, Voorberg J, Reitsma PH, et al. Angiogenic characteristics of blood outgrowth endothelial cells from patients with von Willebrand disease. J Thromb Haemost. 2015;13(10):1854-66.
9. Randi AM, Laffan MA. Von Willebrand factor and angiogenesis: basic and applied issues. J Thromb Haemost. 2017;15(1):13-20.
10. Starke RD, Ferraro F, Paschalaki KE, Dryden NH, McKinnon TA, Sutton RE, et al. Endothelial von Willebrand factor regulates angiogenesis. Blood. 2011;117(3):1071-80.
11. Starke RD, Paschalaki KE, Dyer CE, Harrison-Lavoie KJ, Cutler JA, McKinnon TA, et al. Cellular and molecular basis of von Willebrand disease: studies on blood outgrowth endothelial cells. Blood. 2013;121(14):2773-84.
12. Wang JW, Bouwens EA, Pintao MC, Voorberg J, Safdar H, Valentijn KM, 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):2762-72.
13. de Jong A, Weijers E, Dirven R, de Boer S, Streur J, Eikenboom J. Variability of von Willebrand factor-related parameters in endothelial colony forming cells. J Thromb Haemost. 2019;17(9):1544-54.
14. de Boer S, Bowman M, Notley C, Mo A, Lima P, de Jong A, 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):2721-31.
15. Galili T, O'Callaghan A, Sidi J, Sievert C. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. Bioinformatics. 2017;34(9):1600-2.
16. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nature Genetics. 2000;25(1):25-9.
17. Consortium TGO, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, et al. The Gene Ontology knowledgebase in 2023. Genetics. 2023;224(1).
18. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nature Methods. 2012;9(7):676-82.
19. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A. CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics. 2021;22(1):433.

20. Laan SNJ, Dirven RJ, Bürgisser PE, Eikenboom J, Bierings R. Automated segmentation and quantitative analysis of organelle morphology, localization and content using CellProfiler. PLoS One. 2023;18(6):e0278009.
21. Goncharov NV, Popova PI, Avdonin PP, Kudryavtsev IV, Serebryakova MK, Korf EA, et al. Markers of Endothelial Cells in Normal and Pathological Conditions. Biochem (Mosc) Suppl Ser A Membr Cell Biol. 2020;14(3):167-83.
22. Dejana E, Hirschi KK, Simons M. The molecular basis of endothelial cell plasticity. Nature Communications. 2017;8(1):14361.
23. Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M, et al. Guidelines and definitions for research on epithelial–mesenchymal transition. Nature Reviews Molecular Cell Biology. 2020;21(6):341-52.
24. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nature Protocols. 2019;14(2):482-517.
25. Ferraro F, Kriston-Vizi J, Metcalf DJ, Martin-Martin B, Freeman J, Burden JJ, et al. A two-tier Golgi-based control of organelle size underpins the functional plasticity of endothelial cells. Dev Cell. 2014;29(3):292-304.
26. Medina RJ, O'Neill CL, O'Doherty TM, Wilson SE, Stitt AW. Endothelial progenitors as tools to study vascular disease. Stem Cells Int. 2012;2012:346735.
27. Kat M, Margadant C, Voorberg J, Bierings R. Dispatch and delivery at the ER-Golgi interface: how endothelial cells tune their hemostatic response. FEBS J. 2022;289(22):6863-70.
28. Bowman M, Casey L, Selvam SN, Lima PDA, Rawley O, Hinds M, et al. von Willebrand factor propeptide variants lead to impaired storage and ER retention in patient-derived endothelial colony-forming cells. J Thromb Haemost. 2022;20(7):1599-609.
29. Selvam SN, Casey LJ, Bowman ML, Hawke LG, Longmore AJ, Mewburn J, et al. Abnormal angiogenesis in blood outgrowth endothelial cells derived from von Willebrand disease patients. Blood Coagul Fibrinolysis. 2017;28(7):521-33.
30. Tien TY, Wu YJ, Su CH, Hsieh CL, Wang BJ, Lee YN, et al. Pannexin 1 Modulates Angiogenic Activities of Human Endothelial Colony-Forming Cells Through IGF-1 Mechanism and Is a Marker of Senescence. Arterioscler Thromb Vasc Biol. 2023;43(10):1935-51.
31. Sanchez-Duffhues G, Orlova V, ten Dijke P. In Brief: Endothelial-to-mesenchymal transition. The Journal of Pathology. 2016;238(3):378-80.
32. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119(6):1420-8.
33. Maleszewska M, Moonen JR, Huijkman N, van de Sluis B, Krenning G, Harmsen MC. IL-1β and TGFβ2 synergistically induce endothelial to mesenchymal transition in an NFκB-dependent manner. Immunobiology. 2013;218(4):443-54.
34. Romero LI, Zhang DN, Herron GS, Karasek MA. Interleukin-1 induces major phenotypic changes in human skin microvascular endothelial cells. J Cell Physiol. 1997;173(1):84-92.
35. Kutikhin AG, Tupikin AE, Matveeva VG, Shishkova DK, Antonova LV, Kabilov MR, et al. Human Peripheral Blood-Derived Endothelial Colony-Forming Cells Are Highly Similar to Mature Vascular Endothelial Cells yet Demonstrate a Transitional Transcriptomic Signature. Cells. 2020;9(4).
36. Yoshimatsu Y, Kimuro S, Pauty J, Takagaki K, Nomiyama S, Inagawa A, et al. TGF-beta and TNF-alpha cooperatively induce mesenchymal transition of lymphatic endothelial cells via activation of Activin signals. PLOS ONE. 2020;15(5):e0232356.

37. Hulshoff MS, Xu X, Krenning G, Zeisberg EM. Epigenetic Regulation of Endothelial-to-Mesenchymal Transition in Chronic Heart Disease. Arteriosclerosis, Thrombosis, and Vascular Biology. 2018;38(9):1986-96.
38. Tura O, Skinner EM, Barclay GR, Samuel K, Gallagher RC, Brittan M, et al. Late outgrowth endothelial cells resemble mature endothelial cells and are not derived from bone marrow. Stem Cells. 2013;31(2):338-48.
39. Toshner M, Dunmore BJ, McKinney EF, Southwood M, Caruso P, Upton PD, et al. Transcript analysis reveals a specific HOX signature associated with positional identity of human endothelial cells. PLoS One. 2014;9(3):e91334.
40. Potente M, Mäkinen T. Vascular heterogeneity and specialization in development and disease. Nature Reviews Molecular Cell Biology. 2017;18(8):477-94.
41. Lin Y, Banno K, Gil CH, Myslinski J, Hato T, Shelley WC, et al. Origin, prospective identification, and function of circulating endothelial colony-forming cells in mice and humans. JCI Insight. 2023;8(5).
42. Sánchez-Duffhues G, García de Vinuesa A, van de Pol V, Geerts ME, de Vries MR, Janson SG, et al. Inflammation induces endothelial-to-mesenchymal transition and promotes vascular calcification through downregulation of BMPR2. J Pathol. 2019;247(3):333-46.
43. Rieder F, Kessler SP, West GA, Bhilocha S, de la Motte C, Sadler TM, et al. Inflammation-induced endothelial-to-mesenchymal transition: a novel mechanism of intestinal fibrosis. Am J Pathol. 2011;179(5):2660-73.
44. Derada Troletti C, Fontijn RD, Gowing E, Charabati M, van Het Hof B, Didouh I, et al. Inflammation-induced endothelial to mesenchymal transition promotes brain endothelial cell dysfunction and occurs during multiple sclerosis pathophysiology. Cell Death & Disease. 2019;10(2):45.
45. Medina RJ, O'Neill CL, O'Doherty TM, Chambers SE, Guduric-Fuchs J, Neisen J, et al. Ex vivo expansion of human outgrowth endothelial cells leads to IL-8-mediated replicative senescence and impaired vasoreparative function. Stem Cells. 2013;31(8):1657-68.
46. Smadja DM, Bièche I, Uzan G, Bompais H, Muller L, Boisson-Vidal C, et al. PAR-1 activation on human late endothelial progenitor cells enhances angiogenesis in vitro with upregulation of the SDF-1/CXCR4 system. Arterioscler Thromb Vasc Biol. 2005;25(11):2321-7.
47. Bompais H, Chagraoui J, Canron X, Crisan M, Liu XH, Anjo A, et al. Human endothelial cells derived from circulating progenitors display specific functional properties compared with mature vessel wall endothelial cells. Blood. 2004;103(7):2577-84.
48. Smadja DM, Melero-Martin JM, Eikenboom J, Bowman M, Sabatier F, Randi AM. Standardization of methods to quantify and culture endothelial colony-forming cells derived from peripheral blood. Journal of Thrombosis and Haemostasis. 2019;17(7):1190-4.
49. Blandinières A, Randi AM, Paschalaki KE, Guerin CL, Melero-Martin JM, Smadja DM. Results of an international survey about methods used to isolate human endothelial colony-forming cells: guidance from the SSC on Vascular Biology of the ISTH. J Thromb Haemost. 2023;21(9):2611-9.
50. Smadja DM, Melero-Martin JM, Eikenboom J, Bowman M, Sabatier F, Randi AM. Standardization of methods to quantify and culture endothelial colony-forming cells derived from peripheral blood: Position paper from the International Society on Thrombosis and Haemostasis SSC. J Thromb Haemost. 2019;17(7):1190-4.
51. Cnossen MH, van Moort I, Reitsma SH, de Maat MPM, Schutgens REG, Urbanus RT, et al. SYMPHONY consortium: Orchestrating personalized treatment for patients with bleeding disorders. J Thromb Haemost. 2022;20(9):2001-11.



Supplemental Figure 1. Principal Component analysis (PCA) of bulk RNA transcriptomes. A) Percentage of explained variance is shown in a scree plot. The first 10 principal components are shown. B) Dendrogram of the hierarchical clustering of samples based on PCA 1-4 shows 2 distinct clusters.

Supplemental Table 1. Characteristics of ECFC clones.

Sample number	Donor number	Sex	Age	Day of detection*	Time in culture**	Morph. group	Bulk RNA sequencing	Clustering		Morph. analysis	Migration assay
								RNA seq.	qPCR		
1	1	M	64	13	16	2	Yes	1	1	Yes	No
2	1	M	64	13	18	2	No	-	1	Yes	No
3	1	M	64	13	24	1	No	-	1	Yes	No
4	1	M	64	13	18	1	Yes	1	1	No	No
5	2	M	28	14	26	3	No	-	2	Yes	No
6	3	F	23	15	27	3	Yes	2	2	Yes	No
7	3	F	23	13	23	3	Yes	1	1	Yes	No
8	3	F	23	19	25	2	No	-	2	Yes	No
9	4	F	28	19	30	2	No	-	1	Yes	No
10	5	M	49	15	26	3	No	-	2***	Yes	No
11	6	F	23	28	33	3	Yes	2	2	Yes	No
12	6	F	23	15	19	1	No	-	1	No	No
13	6	F	23	15	14	1	No	-	1	No	No
14	6	F	23	15	19	1	Yes	1	1	Yes	No
15	7	M	29	21	54	2	Yes	2	2	Yes	Yes
16	8	F	26	19	24	2	Yes	1	1	Yes	No
17	9	F	26	14	19	2	Yes	1	1	No	No
18	9	F	26	14	19	2	No	-	1	No	No

Supplemental Table 1. *Continued*

Sample number	Donor number	Sex	Age	Day of detection*	Time in culture**	Morph. group	Bulk RNA sequencing	Clustering		Morph. analysis	Migration assay
								RNA seq.	qPCR		
19	9	F	26	12	18	1	Yes	1	1	Yes	No
20	10	F	24	16	24	3	Yes	2	2	Yes	No
21	10	F	24	19	46	3	No	-	2	Yes	Yes
22	11	M	26	15	15	1	No	-	1	Yes	No
23	11	M	26	15	17	1	Yes	1	1	Yes	Yes
24	12	M	23	16	53	3	No	-	2	Yes	Yes
25	13	M	27	14	23	3	Yes	1	1	Yes	No
26	13	M	27	17	22	2	Yes	1	1	Yes	No
27	13	M	27	14	23	2	No	-	1	Yes	No
28	14	F	20	12	23	1	No	-	1	Yes	No
29	14	F	20	12	21	1	No	-	1	No	No
30	14	F	20	12	15	1	Yes	1	1	Yes	Yes
31	15	F	23	16	40	3	No	-	2	Yes	No
32	16	F	27	13	23	2	No	-	1	Yes	Yes
33	16	F	27	13	21	3	No	-	1	Yes	No
34	10	F	24	15	29	3	No	-	2	Yes	No

*Number of days after inclusion. **From day of detection to freezing (in days). ***Could not be successfully clustered based on available measurements. Cluster 2 was assigned. Abbreviations: Morphological (Morph), Sequencing (seq.)

Supplemental Table 2. Primer sequences used for the qPCR panel

Gene	Category	Sense	Anti-sense
GAPDH	Housekeeping	ACCATCTTCCAGGAGCGAGA	GACTCCACGACGTACTCAGC
RAB27A	WPB content protein	GAAGCCATAGCACTCGCAGAGA	CAGGACTTGTCCACACACCGTT
SELP	WPB content protein	TCCGCTGCATTGACTCTGGACA	CTGAAACGCTCTCAAGGATGGAG
VWF	WPB content protein	TTGACGGGGAGGTGAATGTG	ATGTCTGCTTCAGGACCACG
PLAT	WPB content protein	TGGTGCTACGTCTTTAAGGCGG	GCTGACCCATTCCCAAAGTAGC
CD63	WPB content protein	CAACCACACTGCTTCGATCCTG	GACTCGGTTCTTCGACATGGAAG
ALDH1A1	Proliferation	CGGGAAAAGCAATCTGAAGAGGG	GATGCGGCTATACAACACTGGC
SOX18	Proliferation	GTGTGGGCAAAGGACGAG	GTTCAGCTCCTTCCACGCT
BMP2	Inhibitor of EndMT/EMT	TGTATCGCAGGCACTCAGGTCA	CCACTCGTTTCTGGTAGTTCTTC
COL1A1	EndMT/EMT	CAGCCGCTTCACCTACAGC	TTTTGTATTCAATCACTGTCTTGCC
TGFBi	EndMT/EMT	GGACATGCTCACTATCAACGGG	CTGTGGACACATCAGACTCTGC
IFI27	Apoptosis	CGTCCTCCATAGCAGCCAAGAT	ACCCAATGGAGCCCAGGATGAA
BST2	Apoptosis	TCTCCTGCAACAAGAGCTGACC	TCTCTGCATCCAGGGAAGCCAT
ABI3BP	Senescence	CCTTCTACACCTAAACGACGCC	GGTGTTGTCCATGTAGGTTCAGG
SERPINE1	Inflammation	CTCATCAGCCACTGGAAAGGCA	GACTCGTGAAGTCAGCCTGAAAC
CXCL8	Inflammation	GAGAGTGATTGAGAGTGGACCAC	CACAACCCTCTGCACCCAGTTT

Abbreviations: Weibel-Palade Body (WPB), Endothelial to Mesenchymal transition (EndoMT), Epithelial to Mesenchymal transition (EMT).

Supplemental Table 3. Antibodies used in immunofluorescence

Antibody	Manufacturer	Category number	Concentration	Dilution
IF Primary				
VWF (rabbit)	DAKO	A0082	4.1 mg/mL	1:1,000
VE-cadherin (mouse)	BD Pharming	555661	0.5 mg/mL	1:250
Hoechst	Thermo Fisher Scientific	H3569	10 mg/mL	1:10,000
IF Secondary				
Donkey-anti-Rabbit AF647	Invitrogen Molecular Probes	A31573	2 mg/mL	1:750
Donkey-anti-Mouse AF488	Invitrogen Molecular Probes	A21202	2 mg/mL	1:750

Abbreviations: Von Willebrand factor (VWF), Immunofluorescence (IF)

Supplemental Table 4. Bulk RNA sequencing genes of interest with >7.5 CPM and >2 log2FC

Gene	geneID	Average Log2CPM	Average Log2FC	P-values (FDR corrected)	Category
BST2	ENSG00000130303	7,77	-6,98	9,04E-06	Apoptosis
IFI27	ENSG00000165949	8,45	-4,04	1,48E-02	Apoptosis
IFI6	ENSG00000126709	7,94	-5,07	5,80E-05	Apoptosis
CLU	ENSG00000120885	9,91	4,69	5,80E-05	Apoptosis*
GIMAP4	ENSG00000133574	7,84	2,82	1,22E-04	Apoptosis*
GIMAP6	ENSG00000133561	7,68	2,40	2,70E-04	Apoptosis*
GIMAP8	ENSG00000171115	7,86	2,67	3,00E-04	Apoptosis*
SCARA3	ENSG00000168077	7,55	4,09	3,37E-04	Apoptosis*
ANGPTL4	ENSG00000167772	7,71	-3,42	2,73E-03	EndoMT/ EMT
TGFBI	ENSG00000120708	8,63	-4,45	1,21E-03	EndoMT/ EMT
ALCAM	ENSG00000170017	8,51	-2,43	2,53E-04	EndoMT/ EMT
CADM3	ENSG00000162706	8,10	-6,60	4,71E-04	EndoMT/ EMT
CD44	ENSG00000026508	7,59	-2,93	4,71E-04	EndoMT/ EMT
COL1A2	ENSG00000164692	9,24	-5,22	1,59E-03	EndoMT/ EMT

Supplemental Table 4. *Continued*

Gene	geneID	Average Log2CPM	Average Log2FC	P-values (FDR corrected)	Category
COL5A1	ENSG00000130635	10,00	-2,72	1,13E-03	EndoMT/ EMT
COL5A2	ENSG00000204262	9,77	-2,09	1,27E-03	EndoMT/ EMT
COL8A1	ENSG00000144810	9,86	-2,14	4,88E-03	EndoMT/ EMT
FN1	ENSG00000115414	12,95	-2,07	2,13E-03	EndoMT/ EMT
ITGAV	ENSG00000138448	9,73	-2,87	7,20E-04	EndoMT/ EMT
LTBP1	ENSG00000049323	8,64	-2,58	1,13E-03	EndoMT/ EMT
NRCAM	ENSG00000091129	7,55	-2,67	2,35E-03	EndoMT/ EMT
PLOD2	ENSG00000152952	8,63	-2,18	5,76E-04	EndoMT/ EMT
PVR	ENSG00000073008	8,44	-2,65	1,22E-03	EndoMT/ EMT
TPM1	ENSG00000140416	9,93	-2,48	3,30E-04	EndoMT/ EMT
TPM2	ENSG00000198467	7,76	-2,41	9,64E-03	EndoMT/ EMT
MMRN1	ENSG00000138722	11,20	5,18	9,45E-05	Endothelial marker
TFPI	ENSG00000003436	8,46	2,18	1,19E-03	Endothelial marker
TSPAN18	ENSG00000157570	7,57	2,78	1,82E-04	Endothelial marker
VWF	ENSG00000110799	12,67	2,63	5,04E-04	Endothelial marker
CXCL1	ENSG00000163739	8,65	-3,72	6,97E-03	Inflammation
CXCL6	ENSG00000124875	9,28	-6,10	5,18E-03	Inflammation
CXCL8	ENSG00000169429	8,81	-3,67	1,15E-02	Inflammation
IL1RL1	ENSG00000115602	7,64	-2,86	2,29E-03	Inflammation
IL32	ENSG00000008517	7,58	-2,09	9,14E-03	Inflammation
PTX3	ENSG00000163661	10,89	-2,72	3,98E-03	Inflammation
SERPINE1	ENSG00000106366	12,90	-2,62	7,28E-03	Inflammation
SLC7A2	ENSG00000003989	7,89	-3,86	2,56E-04	Inflammation
MYH10	ENSG00000133026	8,87	2,24	2,72E-04	Migration
EMCN	ENSG00000164035	8,12	2,11	1,32E-02	Proliferation
ADGRG6	ENSG00000112414	7,54	2,10	2,08E-02	Proliferation
ALDH1A1	ENSG00000165092	9,54	7,20	5,64E-04	Proliferation

Supplemental Table 4. *Continued*

Gene	geneID	Average Log2CPM	Average Log2FC	P-values (FDR corrected)	Category
ELK3	ENSG00000111145	8,32	2,55	5,76E-04	Proliferation
NDRG4	ENSG00000103034	7,57	2,63	2,40E-03	Proliferation
PPP1R16B	ENSG00000101445	8,13	2,29	1,66E-04	Proliferation
CAVIN2	ENSG00000168497	9,28	2,85	3,45E-05	Proliferation/ migration
DCBLD2	ENSG00000057019	7,87	-2,23	4,69E-03	Proliferation/ migration*
PDE2A	ENSG00000186642	7,89	3,40	9,48E-04	Proliferation/ migration
SOX18	ENSG00000203883	7,99	3,35	2,49E-04	Proliferation/ migration
TAGLN	ENSG00000149591	7,95	-3,02	3,73E-02	Senescence
ABI3BP	ENSG00000154175	9,24	-4,49	3,30E-03	Senescence
SOD2	ENSG00000112096	9,66	-2,96	7,67E-03	†
ABLIM1	ENSG00000099204	7,93	2,44	2,15E-03	†
ESM1	ENSG00000164283	8,31	-2,13	2,32E-02	†
HLA,B	ENSG00000234745	9,61	-2,29	4,35E-04	†
DIPK2B	ENSG00000147113	7,71	2,10	4,34E-03	†

Abbreviations: Count Per Million (CPM), Fold Change (FC), False Discovery Rate (FDR), Endothelial/ epithelial to Mesenchymal Transition (EndoMT/EMT) Symbols: *negatively regulates or inhibits this process, †could not be placed in a category