UNIVERSITY OF BIRMINGHAM

University of Birmingham Research at Birmingham

Proteome Changes during Transition from Human Embryonic to Vascular Progenitor Cells

Tsolis, Konstantinos; Bagli, Eleni; Kanaki, katerina; Zografou, Sofia; Carpentier, Sebastien; Bei, Ekaterini; Christoforidis, Savvas; Zervakis, Michalis; Murphy, Carol; Fotsis, Theodore; Economou, Anastassios

Document Version
Peer reviewed version

Link to publication on Research at Birmingham portal

Proteome Changes during Transition from Human Embryonic to Vascular Progenitor Cells

Tsolis, Konstantinos; Bagli, Eleni; Kanaki, katerina; Zografou, Sofia; Carpentier, Sebastien; Bei, Ekaterini; Christoforidis, Savvas; Zervakis, Michalis; Murphy, Carol; Fotsis, Theodore; Economou, Anastassios

Link to publication on Research at Birmingham portal

Proteome changes during transition from human embryonic to vascular progenitor cells

Konstantinos C. Tsolis[1,2,4!], Eleni Bagli[3!], Katerina Kanaki[2!], Sofia Zografou[3], Sebastien Carpentier[5,6], Ekaterini S. Bei[7], Savvas Christoforidis[3,8], Michalis Zervakis[7], Carol Murphy[3,9], Theodore Fotsis[3,8,9], and Anastassios Economou[1,2,4,6*]

[1]KU Leuven - University of Leuven, Department of Microbiology and Immunology, Rega Institute for Medical Research, Laboratory of Molecular Bacteriology, B-3000 Leuven, Belgium.

[2]Institute of Molecular Biology and Biotechnology - FORTH, Department of Protein structure and Proteomics Facility, Iraklio, Crete.

[3]Institute of Molecular Biology and Biotechnology - FORTH, Division of Biomedical Research, 45110 Ioaninna, Greece.

[4]Department of Biology, University of Crete, Iraklio, Crete, Greece

[5]KU Leuven - University of Leuven, Department of Biosystems Engineering, B-3000 Leuven, Belgium.

[6]SYBIOMA, KU Leuven facility for Systems Biology based Mass spectrometry, Leuven Belgium

[7]School of Electronic and Computer Engineering, Technical Univ. of Crete, Chania, Greece.

[8]Laboratory of Biological Chemistry, Medical School, University of Ioannina, 45110 Ioannina, Greece.

[9]School of Biosciences, College of Life and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK.

[!]These authors contributed equally

[*]for   correspondance: tassos.economou@rega.kuleuven.be

e-mail addresses

Konstantinos C Tsolis	E-mail: konstantinos.tsolis@rega.kuleuven.be
Eleni Bagli	E-mail: elenibgl@hotmail.com
Katerina Kanaki	E-mail: katkanaki@hotmail.com
Sofia Zografou	E-mail: szografu@cc.uoi.gr
Sebastien Carpentier	E-mail: sebastien.carpentier@biw.kuleuven.be
Ekaterini S Bei	E-mail: beieka@yahoo.gr
Savvas Christoforidis	E-mail: schristo@uoi.gr
Michalis Zervakis	E-mail: michalis@display.tuc.gr
Carol Murphy	E-mail: murphyce@adf.bham.ac.uk
Theodore Fotsis	E-mail: thfotsis@uoi.gr
Anastassios Economou	E-mail: tassos.economou@rega.kuleuven.be
	phone: +32 16 37 92 73

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abbreviations

ACD:	Asymmetric Cell Division
BH:	Benjamini & Hochberg
BMP:	Bone morphogenic protein
EPCs:	Endothelial progenitor cells
ERM:	Ezrin, Radixin, Moesin protein family
FBS:	Fetal bovine serum
FWHM:	Full width half maximum
GMP:	Good Manufacturing Practice
GSK:	Glycogen Synthase Kinase
hESCs:	Human embryonic stem cells
LFQ:	Label Free Quantification
IF:	Immunofluoresence
iPSCs:	Induced Pluripotent Stem Cells
PC:	Principal Component
PCs:	Pericytes
PDGF:	Platelet-derived growth factor
PLS:	Partial Least Squares regression
PMSF:	Phenyl-methyl-sulfonyl Fluoride
RM:	Regenerative Medicine
SDS:	Sodium Dodecyl Sulfate
SMCs:	Smooth muscle cells
VEGF:	Vascular Endothelial Growth factor
VIP:	Variable Importance in Projection
VPCs:	Vascular Progenitor Cells
VSMCs:	Vascular smooth muscle cells

Keywords: Endothelial progenitor cells, hESCs, vascular cell differentiation, CD34[+], angiogenesis, comparative proteomics, multivariate analysis, PLS.

Abstract

Human embryonic stem cells (hESCs) are promising in Regenerative Medicine (RM), due to their differentiation plasticity and proliferation potential. However, a major challenge in RM is the generation of a vascular system, to support nutrient flow to newly synthesized tissues. Here we refined an existing method to generate tight vessels, by differentiating hESCs in $CD34^+$ Vascular Progenitor Cells (VPCs), using chemically defined media and growth conditions. We selectively purified these cells from $CD34^-$ outgrowth populations also formed. To analyze these differentiation processes, we compared the proteomes of the hESCs with those of the $CD34^+$ and $CD34^-$ populations, using high resolution mass spectrometry, label-free quantification and multivariate analysis. 18 protein markers validate the differentiated phenotypes in immunological assays; 9 of these were also detected by proteomics and show statistically significant differential abundance. Another 225 proteins show differential abundance between the three cell types. 63 of these have known functions in $CD34^+$ and $CD34^-$ cells. $CD34^+$ cells synthesize proteins implicated in endothelial cell differentiation and smooth muscle formation, supporting the bipotent phenotype of these progenitor cells. $CD34^-$ cells are more heterogeneous synthesizing muscular/osteogenic/chondrogenic/adipogenic lineage markers. The remaining >150 differentially abundant proteins in $CD34^+$ or $CD34^-$ cells raise testable hypotheses for future studies to probe vascular morphogenesis.

Data are available via ProteomeXchange with identifier PXD003606.

Introduction

Regenerative Medicine (RM) is an emerging interdisciplinary field of research and clinical applications, focused on the repair, replacement, or regeneration of cells, tissues, or organs to restore impaired function resulting from any cause, including congenital defects, disease, and trauma.[1] Human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs), created by reprogramming adult donor cells, have the ability to differentiate into any human cell and ultimately to generate any human tissue.[2-4] However, main challenges in RM are the ability to reproducibly differentiate hESCs/iPCSs to a specific cell type and the creation of vascular tissue to ensure rapid *in vivo* vascularization and sufficient nutrient flow to the implant inside the host.[5-8] In addition, RM protocols require compliance with the GMP (Good Manufacturing Practice) specifications necessitating chemically defined media to avoid use of serum or feeder layers, ensuring that the cells have been produced according to predefined manufacturing criteria.[9]

Current interest has shifted towards generating tissue-engineered constructs that are already vascularized before implantation.[10] For the differentiation process, cells are grown on biodegradable 3D scaffolds, which mimic cell-cell or cell-matrix interactions. In addition, soluble factors are mimicking the physiological microenvironment leading to differentiation of hESCs to Vascular Progenitor Cells (VPCs)**.** Differentiation of VPCs from their precursor cells involves co-operative interaction among many different signaling molecules [i.e. hedgehog, vascular endothelial growth factor (VEGF), basic fibroblast growth factor (bFGF), bone morphogenetic protein (BMP), Notch, Wnt], and transcription factors (ETS, Forkhead, GATA transcription factors).[11, 12] A thorough understanding of the molecular events that underlie this differentiation process is still missing as are comprehensive biomarker signatures with which to monitor both the transition as well as the biological robustness of the final differentiated states.[13]

Omics workflows provide powerful means for the study of the molecular base of the differentiation process of VPCs, due to their ability to monitor multiple targets (genes/proteins) simultaneously, in complex samples. Several RNA-based omics approaches have been used for the study of vascular progenitor cell differentiation.[14-19] However, mRNA expression levels are not necessarily correlated with protein abundance.[20] Hence proteomics approaches are essential to directly analyze the expressed proteome shifts that occur during differentiation events and the concomitant post-translational and complexome changes.[21] To our knowledge, there are no proteomics analyses in the literature of VPCs derived from human pluripotent SCs, in defined culture conditions. Several analyses have been performed in ESCs or iPSCs[22], in primary endothelial cells after stimulation with VEGF[23], or in *in vitro* differentiated erythroid cells derived from iPSCs.[23-25] Endothelial cells display functional heterogeneity that cannot be addressed with the existing markers for cell classification.[13] Therefore, there is intense need to identify more VPCs markers[26, 27], which can be used to monitor this population. Proteomics can be a major contributor in the development of proteomic signatures of VPCs.

Here, we modified and refined an existing method in order to generate VPCs that could be used to generate tight vessels. hESCs were first differentiated to mesodermal intermediates and then towards Vascular Progenitor cells (VPCs). Each growth factor is added separately, for better control over the differentiation process and offers the possibility to study each step separately. In addition, the protocol is fast, providing functional VPCs within 5 days. The differentiation event gives rise to two distinct cellular populations: the $CD34^+$ VPCs and a $CD34^-$ population that probably contains mixed cells (see below), that can be further differentiated into other cell types such as other mesoderm derivatives including cells from the mesenchymal lineage.[28] We monitored these cell states by comparing the proteome of the initial hESCs with the differentiated $CD34^+$ and $CD34^-$ populations, using high resolution mass spectrometry. 236 proteins with differential abundance were identified. 46, 86 and 51 proteins

are over-synthesized in hESCs, in CD34$^+$ and CD34$^-$, respectively. 18 known protein markers of the differentiation cell state were detected using immunological assays, and 9 of them are also detected in the proteomics flow and show statistically significant differential abundance. Some of the identified proteins have known functions in CD34$^+$ or CD34$^-$ cells based on the literature and thus provide additional validation for our differentiation protocol and the proteomics flow. Another ~150 proteins are new candidates for further hypothesis-driven testing regarding the phenotype of CD34$^+$ VPCs or CD34$^-$ cells and the mechanism of vascular cell differentiation.

Experimental Section

hESC culture in feeder-free conditions

The H1 hESC line was purchased from Wicell Research Institute (Madison, WI, USA) and maintained on six-well tissue culture plates coated with hESC-qualified Matrigel (Corning, Cat.No.354277) in complete mTeSR[TM]1 medium (05850, Stem Cell Technologies). Cells were routinely characterized and found to be karyotypically normal (data not shown) and express the pluripotent markers OCT4, SSEA4, SOX2, NANOG and alkaline phosphatase as determined by western immunostaining (Supporting Figure S-1, S-2).[29] Every 5-7 days, cells were passaged enzymatically using 1 mg/ml dispase (Invitrogen) for 2 minutes at 37°C. hESCs colonies were then harvested, dissociated into small clumps and re-plated onto Matrigel-coated 6-well plates (ratio 1:6).

Differentiation of hESCs

Differentiation of hESCs to CD34[+] cells was carried out under feeder-free, chemically defined conditions as described with minor modifications (Fig. 1A).[28] Briefly, H1 colonies were first dissociated into small clumps and re-plated onto Matrigel-coated 6-well plates as per normal routine passaging. After 48 h, mTeSR medium was changed to differentiation medium (APEL), which was synthesized as previously described [30], supplemented with 5 μM Glycogen Synthase Kinase-3 (GSK-3) inhibitor (CHIR99021, Selleckchem). After 24 h, the medium was replaced with the same basal medium supplemented with Bone Morphogenetic Protein-4 (BMP-4) (25 ng/ml; PHC9534, Life Technologies) for 48 h and then with Vascular Endothelial Growth Factor-A (VEGF-A) (80ng/ml; Immunotools) for another 48 h. On day 5, successful commitment to endothelial lineage was confirmed by the presence of a CD34[+]KDR[+] cell population using FACs analysis as previously described (Supporting Table S-1).[28]

Isolation of CD34$^+$ cells

On day 5 of the differentiation process, CD34$^+$ cells were isolated using the EasySep™ Human CD34 Positive Selection Kit (Stem Cell Technologies, Cat.No.18056) according to the manufacturer's instructions (see Supporting Experimental Procedures). In total, we performed 6 biological repeats of the differentiation experiment were we studied hESCs, CD34$^+$ and CD34$^-$ populations separately.

Protein extraction, sample preparation for proteomic analysis

Total cells were lysed using RIPA lysis buffer (50 mM Tris/HCl pH 7.2, 150 mM NaCl, 1 % v/v Triton X-100, 0.5 % w/v sodium deoxycholate, 0.1% w/v SDS, 2.5 mM PMSF) (see Supporting Experimental Procedures). 40 μg total protein extract from each repeat was analysed in 1D-SDS-PAGE. Gels were stained with colloidal coomassie blue (0.12% w/v Coomassie G250, 10% v/v phosphoric acid, 10% w/v ammonium sulfate, 20% v/v methanol) and each lane was cut into 10 slices. Each slice was trypsin digested and the corresponding peptides were stored until the LC-MS/MS analysis (see Supporting Experimental Procedures).

LC-MS/MS analysis

Lyophilized peptide samples were first dissolved in an aqueous solution containing 0.1% v/v formic acid (FA) and 5% v/v ACN and afterwards were analyzed using nano-Reverse Phase LC coupled to a Q Exactive™ Hybrid Quadrupole - Orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany) through a nanoelectrospray ion source (Thermo Scientific, Bremen, Germany). Peptides were initially separated using a Dionex UltiMate 3000 UHPLC system on an EasySpray C18 column (Thermo Scientific, OD 360 μm, ID 50 μm, 15 cm length, C18 resin, 2 μm bead size) at a nanoLC flow rate of 300 nL min^{-1}. The LC mobile phase consisted of two different buffer solutions, an aqueous solution containing 0.1% v/v FA (Buffer A) and an aqueous solution containing 0.08% v/v FA and 80% v/v ACN (Buffer B). A 60 min

multi-step gradient was used from Buffer A to Buffer B (percentages from each in parentheses below) as follows: 0–3 min constant (96:4), 3–15 min (90:10); 15–35 min (65:35); 35–40 min (35:65); 40-41 min (5:95); 41-50 min (5:95); 50-51 min (95:5); 51-60 min (95:5).

The separated peptides were analyzed in the Orbitrap QE operated in positive ion mode (nanospray voltage 1.5 kV, source temperature 250°C). The instrument was operated in data-dependent acquisition (DDA) mode with a survey MS scan at a resolution of 70,000 FWHM for the mass range of *m/z* 400-1600 for precursor ions, followed by MS/MS scans of the top 10 most intense peaks with +2, +3 and +4 charged ions above a threshold ion count of 16,000 at 35,000 resolution. MS/MS was performed using normalized collision energy (NCE) of 25% with an isolation window of 3.0 *m/z*, an apex trigger 5-15 sec and a dynamic exclusion of 10 s. Data were acquired with Xcalibur 2.2 software (Thermo Scientific).

MS data analysis

Raw MS files from the mass spectrometer were analyzed by MaxQuant v1.5.2.8, a quantitative proteomics software package designed for analyzing large mass spectrometric data sets.[31] MS/MS spectra were searched by the Andromeda search engine against the Uniprot human reviewed proteome without isoforms (Last modified - July 6, 2015, 20,198 proteins) and common contaminants as described.[32] Enzyme specificity was set to trypsin, allowing for a maximum of two missed cleavages. Dynamic (methionine oxidation and N-terminal acetylation) and fixed (S-Carbamidomethylation of cysteinyl residues) modifications were selected. Precursor and MS/MS mass tolerance was set to 20 ppm for the first search (for the identification of maximum number of peptides for mass and retention time calibration) and 4.5 ppm for the main search (for the refinement of the identifications). Protein and peptide False Discovery Rate (FDR) were set to 1%. Peptide features were aligned between different runs and masses were matched ("match between runs" feature), with a match time window of 0.7 min and a mass alignment window of 10 min. Proteins were normalized and quantified using the

MaxLFQ algorithm[33] with a minimum of 2 peptides per protein (18.752 or 27% of the sum of proteins identified in all samples was excluded). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD003606.[34]

Multivariate statistical Analyses

For the identification of differentially synthesized proteins we performed multivariate analysis using the Partial Least Square regression (PLS) method (n = 6).[35, 36] We constructed a PLS model using the LFQ intensities of the whole dataset, derived from the three populations (hESCs, CD34$^+$ and CD34$^-$ cells). Most significant variables were selected after performing Variable Importance in Projection analysis (VIP).[37] For the non-identified proteins, the quantitative value was set to 0. The 250 most significant proteins, based on the VIP analysis, were then tested for statistically significant difference in abundance. Since our data does not necessarily follow the assumptions of normally distributed data of equal variance, we performed the non-parametric test Kruskal-Wallis (p-value <0.05), on LFQ intensities to select proteins of possible interest. Pairwise differences were then examined using the Wilcoxon test (p-value <0.05).

Other software

Images and vectors were processed using Canvas (ACDSee). Statistical analysis was performed using STATISTICA v8 (Statsoft) or R scripts, and graphs were plotted in R using ggplot2 package or in Prism v5 (GraphPad).

Results

Differentiation of hESCs to CD34+ Vascular Progenitor Cells

Differentiation of hESCs to vascular progenitor cells was conducted in a chemically defined stepwise approach, using a modified protocol from Tan et al.[28] Pluripotent stem cells were first induced to mesoderm using a GSK inhibitor and BMP4 and then to vascular commitment with an additional VEGF treatment (48 h; Fig. 1A). On the 5th day of the differentiation procedure small cobblestone-like cells were spreading outwards from the edges of the colony (Supporting Fig. S-1A). FACs analysis revealed that 25-40% of the cells co-express the CD34, CD31 and KDR surface markers (Fig. 1B & C, Supporting Fig. S-1B), which are not expressed in hESCS (42-44), being used for the identification of vascular progenitor cells. Without VEGF treatment, CD34 and CD31 remained down-regulated (data not shown), indicating that stimulation with VEGF is an essential step for the differentiation process. Since CD34 has generally been associated with progenitors of endothelial lineages, we isolated the CD34+ cells using magnetic beads coated with anti-CD34 antibody (Fig. 1C).

First, we assessed the endothelial potential of the isolated hESC-derived CD34+ cells by testing them for tube formation on 3D Matrigel cultures. Indeed, the sorted CD34+ fraction (but not the CD34- fraction) could be organized into vascular-like structures on a Matrigel support (Fig. 1D). Quantitative RT-PCR analyses showed that CD34+ cells expressed typical endothelial markers such as the genes encoding VEGF receptors (VEGFR 1,2,3; neuropillin 1, 2), CD31 (*PECAM1*), VE-Cadherin (*CDH5*), Tie2 (*TEK*)[38], the arterial endothelial marker EPHRINB2 (*EFNB2*)[39] and the venous endothelial marker EPHRINB4 (*EFNB4*)[40] (Supporting Fig. S-1F). Also *ETV2*, a transcription factor belonging to the ets-transcription family, which is important for the commitment to the endothelial lineage, during development,[41-45] was upregulated (Supporting Fig. S-1F). However, the gene for von Willebrand factor (*vWF*), a marker of mature

endothelial cells[38], showed marginal expression demonstrating the premature nature of the purified CD34[+] cells. Finally, the expression of genes typical to mural cells[46, 47] such as calponin (*CNN1*), smooth muscle protein 22-alpha (*SM22/TAGLN*) and Chondroitin Sulfate Proteoglycan NG2 (*NG2/CSPG4*) was negligible (Supporting Fig. S-1F).

To investigate the differentiation potential of CD34[+] progenitor cells to mature vascular cells (endothelial and mural cells), post-sorted CD34[+] cells were cultured on fibronectin coated dishes in APEL differentiation medium supplemented with either VEGF (50 ng/ml) or PDGFB (50 ng/ml) and TGFβ1 (5 ng/ml). Cells treated with VEGF exhibited cobblestone morphology, were positive for the endothelial marker CD31 and did not express the contractile mural marker SM22-alpha demonstrating their endothelial phenotype (Supporting Fig. S-2). In contrast, when the selected CD34[+] cells were cultured with PDGFβ/TGFβ, exhibited a fibroblast-like morphology, lost the expression of CD31 and expressed the typical mural marker SM22-alpha (Supporting Fig. S-2B). These data validate that VPCs produced with the described protocol, have the capacity to further differentiate into mature vascular cells.

Proteomic analysis of pluripotent hESCs, CD34[+] and CD34[-] cells

We first compared the proteomic profile of hESCs with that of the CD34[+] and CD34[-] populations (see Experimental Section). To exclude proteins with less reliable quantification, from the total number of proteins identified we proceeded for further analysis using the ones that were quantified using two peptides. In total, 4491 proteins were quantified, in all groups, at least once (Supporting Table S-2). Label-free quantification reproducibility between the biological repeats, is good for the specific workflow setup that was followed, with average Pearson correlation $r^2 = 0.86$ between biological repeats of the same group and $r^2 = 0.83$, between groups (Supporting Fig. S-3A). Protein copy numbers in the cell span roughly in a range of seven orders of magnitude.[48] Given the similarity between different cell types[49, 50], we approximated the abundance of the proteins identified here based on a reference set by aligning

the experimentally identified proteins of this study with those of the reference one (Supporting Fig. S-3B).[48] Proteins present in low copy numbers are also identified here, suggesting good identification depth.

Next, we sought out differentially synthesized proteins by performing multivariate analysis. Due to the nature of the proteomics data (small number of samples (n) and large number of variables – proteins (p)), we implemented Partial Least Squares regression (PLS) to identify possible proteins of interest[35, 51], using the LFQ-derived abundance values of the identified proteins. To select the proteins that show the greatest variation in our dataset, we combined PLS analysis with the Variable Importance in Projection (PLS-VIP) method.[36, 37] By plotting the scores of the PLS model, we observe 3 distinct clusters corresponding to the 6 biological repeats of each of the three different cell populations examined (Fig. 2A). Also, the variation between the biological repeats is smaller in hESCs and CD34[+] cells types compared to that in the CD34[-] cells, which are considered to be a mixed population of several other progenitors. To find proteins with differential abundance, we then used the VIP method to identify the proteins with the highest variation. The 250 proteins that contribute most to the variation were then tested for statistical significance. Since we cannot consider that the proteomics quantification approach that we use, follows the assumptions of normality or equal variance between the different study groups, we used the non-parametric test Kruskal-Wallis (p-value < 0.05) followed by *post hoc* analysis using the Wilcoxon test (p-value < 0.05). 236 proteins showed differential synthesis in at least one pair (Fig. 2B & C, Supporting Table S-3, Supporting Fig. S-4).

Validation of proteomics results using established protein markers

To validate the quantification approach of our proteomics workflow, we assessed the expression of selected markers in hESCs and in their derived differentiated populations, using immunofluorescence (IF), western blot analysis (WB) and flow cytometry (FACS). Loss of pluripotency during this differentiation procedure was evidenced by the down-regulation of

NANOG, SOX2, OCT3/4 (POU5F1), E-Cadherin (CDH1) and Alkaline phosphatase (ALPL) (Supporting Fig. S-2A) that are the typical pluripotent stem cell markers.[52-54] Typical surface endothelial markers KDR, VE-Cadherin (CDH5), CD31, Endothelial protein C receptor PROCR (CD201) as well as the endothelial transcription factor ERG revealed robust and almost exclusive expression in CD34[+] cells by FACs and western immunoblots (Fig. 1E; Supporting Fig. S-1C).[55] Surface expression of the characteristic vascular markers, CD73, CD44, CD105, CD146 (MCAM) and CD140B (PDGFRB) known to be expressed in endothelial and mural cells, was evaluated using FACs.[40, 46, 56, 57] As expected, none of these markers was expressed in hESCs (<2% positive; Supporting Fig. S-1A). CD73 was expressed only in the CD34[+] population and the remaining markers were differentially expressed among the CD34[+] and CD34[-] cell fractions. Specifically, there was a clear shift in the intensity for the surface expression of CD44, CD105 and CD146 in the whole CD34[+] population compared to that in CD34[-], in FACs analysis, which could be possibly interpreted as a higher total expression of these markers in the CD34[+] fraction. Analysis of CD140B showed a double population, where 43 ± 5,5 % of CD34[+] and 66 ± 4,5 % of CD34[-] cells, were positive for the CD140B marker. Among the 18 markers validated with immunological methods, 9 of them (OCT4 (POU5F1), ALPL, CD34, KDR, CDH5 (VE-Cadherin), PROCR (CD201), ERG, CD44 and PDGFRB) display the same profile in the proteomics and the immunological methods and show statistically significant differential abundance between the study groups (Fig. 4A), validating our proteomics flow.

GO-slim analysis of the differential proteins

To gain functional insight in the proteins identified by the proteomics flow, we performed limited scale Gene Ontology analysis (GO-slim)[58] for the differential proteins in the three cell populations (hESCs, CD34[+] VPCs and CD34[-]). Proteins specific to each group (see Experimental Section), were uploaded separately in the AmiGO server and tested for selected

GO terms (Fig. 3). hESCs are enriched in proteins related to the chromosome and methylation processes (Supporting Table S-4). Proteins enriched in CD34[+] cells are grouped into three categories: a. vesicle-mediated transport or localized in membrane-bound vesicles; b. nitric oxide metabolic processes, that is in agreement with the endothelial phenotype of this population[59, 60]; and c. proteins localized in the cytoskeleton or are related with biological adhesion and locomotion. On the other hand, in CD34[-] cells we identify more proteins related with biological adhesion and locomotion, compared to hESCs, however to a lower extent than in the CD34[+] cells. Overall, GO-slim analysis, suggests that the characteristic phenotypic shift from that of the hESCs to that of the CD34[+] or CD34[-] populations, is underscored by two main processes, adhesion/migration and vesicle-mediated transport.

Pathway Enrichment Analysis

We then sought to determine differentially activated pathways between hESCs and CD34[+]/CD34[-] cells in our experimental dataset. For this we used WebGestalt, an *in silico* tool for functional annotation, which incorporates information from different public resources for the biological interpretation of the omics data.[61, 62] Differentially synthesized proteins from hESCs, CD34[+] and CD34[-] cells were uploaded as one dataset in WebGestalt (Supporting Table S-3). Enrichment was performed in pathways included in the publicly curated database Wikipathways (see Supporting Experimental Procedures). Annotation enrichment could identify fifteen pathway terms that were statistically over-represented within the list of the differential proteins from all three cell populations (Table 1). These pathways represent: i) adhesion and mechano-transduction, ii) signaling and iii) regulation or metabolism.

Manual curation of the proteomics results – Cell characterization

As a next step to improve our understanding about the phenotype of the three cell populations analyzed here, we checked manually for references of the differentially synthesized

proteins in the literature. Since hESCs are well described, we focused more in the CD34$^+$ and CD34$^-$ cells.

Our proteomic analysis revealed that 84 proteins were significantly more abundant in the isolated CD34$^+$ cells (Supporting Table S-4). Typical vascular markers such as KDR, PROCR, NOS3, CD44, CDH5 are included in that list, as was expected. The vascular progenitor status of the isolated CD34$^+$ cells is also supported by the fact that of the 84 up-regulated proteins we identified, 33 are known to be essential for vasculogenesis/angiogenesis, vascular homeostasis and endothelial function in general (Supporting Table S-5, Fig. 4B). Isolated CD34$^+$ cells showed elevated synthesis of 7 proteins that are expressed in immature smooth muscle cells and play a role in vascular smooth muscle differentiation and function (PALD1, HDAC7, TLN1, VASP, CRIP2, LMNA, CORO1C) (see references in Supporting Table S-6; Fig. 4C), suggesting that CD34$^+$ cells exhibit a bipotent phenotype that can differentiate to both mural (vascular muscle cells and pericytes) and endothelial cells depending on the conditions (Fig. 5,I and II). Indeed, treatment of CD34$^+$ cells with PDGF generates cells that express the sm22 protein, characteristic for smooth muscle cells (Supporting Fig. S-2). Support for the bipotent phenotype of CD34$^+$ cells derives from the expression pattern of the PALD1 protein, which during vascular development is expressed in endothelial cells of the embryo, but in adulthood its expression is confined selectively to arterial smooth muscle cells.[63]

Discussion

Cell-based vascular engineering to regenerate and remodel damaged or structurally abnormal vessels, and the affected tissues thereof, constitutes an important area of research requiring further basic research before clinical translation. Regenerating tissues over 100–200 µm exceeds the capacity of nutrient supply and waste removal by diffusion, thus requiring an intimate supply of vascular networks[7, 8] to ensure survival of the implant *in vivo*.[64] Pre-vascularized tissue-engineered constructs (TECs) consisting of mature vessels should be able to connect in a very short time to the vessels of the host providing immediate blood supply to the TEC. Despite the abundance of preclinical animal studies[65-67], there are still incomplete data and a lack of concrete guidelines concerning the generation of pre-vascularized TECs for preclinical evaluation. However, the current use of vascular cells differentiated from hESC/hiPSC are superior to the adult stem cell-derived VPCs. This is due to the higher proliferative potential and plasticity, because scaling TECs to human size requires maintaining constant physical conditions and cellularity over larger dimensions.

An important issue in vessel regeneration is the initial source of the cells to be used in therapeutic approaches. Generation of engineered vessels was achieved using a vast spectrum of endothelial cells of different differentiation status such as HUVECs[68], cord blood-derived progenitors[69] or hESCs[70] co-implanted with murine MPC 10T1/ cells[68-70] or human MSCs[71] indicating that a functional stable vasculature for tissue engineering requires both ECs and a source of mural cells (vascular smooth muscle, MSC, pericyte, or MPC.[72] Moreover, it is also feasible to obtain mural/perivascular cells from hESCs/hiPS cells.[12, 73] A critical issue in the entire approach is the accurate identification of sub-populations of progenitors that have specific lineage commitments. Currently, the isolation of vascular progenitor cells relies primarily on the expression of the hematovascular stem marker CD34[+] and the co-expression of a limited number of general surface markers such as KDR, CD31, CD144 (VE-cadherin).[65-67] Since

CD34[+] cells represent a progenitor cell population with therapeutic potential, the detailed characterization of its dynamic phenotype becomes crucial. While RNA-based approaches have been important in delineating the roles of specific genes in vascular progenitor cell characterization[14-19], proteomics studies help provide a more direct and representative picture of the actual differentiation state of the cells.

Here we modified and refined an existing protocol, for the differentiation of hESCs into Vascular Progenitor Cells (VPCs) (CD34[+] cells)[28], in defined media, under feeder free conditions. After 5 days of culture, two main cell types are generated, the CD34[+] VPCs and a CD34[-] population. The differentiation phenotype is stable and reproducible, as shown by immunophenotyping of 18 known protein markers (Supporting Fig. S-1). Isolated CD34[+] cells show functional characteristics of vascular progenitor cells in an *in vitro* tube formation assay.

Differential protein abundance from the initial hESCs to the derived CD34[+] and CD34[-] populations was monitored using high resolution MS combined with multivariate analysis of 6 biological repeats. PLS regression identified proteins showing the highest variation between the biological repeats, using the LFQ abundance values. Due to the small number of samples (biological repeats) over the number of variables (identified proteins), PLS provides a more accurate model over the traditional regression methods.[51] The 250 proteins showing the most variation between the groups, were selected using the VIP method[37], and further validated using the non-parametric test Kruscal-Wallis. Nine of the markers tested during immunophenotyping were also detected by proteomics in the differentially abundant proteins, thus validating our proteomics flow. Among the possible proteins of interest, 63 present in CD34[+] or CD34[-] could be assigned to known biological functions related to vasculogenesis, providing a broad view of this differentiation event and enhancing previous hypothesis. In addition, we identify >150 proteins showing differential abundance that might help in the understanding of the mechanism of vasculogenesis, or could provide possible differentiation markers (Supporting Table S-4).

The isolated CD34$^+$ population is apparently a vascular progenitor stem cell population with commitment to an endothelial cell lineage. Not only it expresses all the known markers associated with this status (KDR, FLT1, FLT4, NRP1 and 2, TEK, EFNB2, EPHB4, PECAM1, CDH5, ETV2, HIF1A, vWF), but also the great majority of the identified proteins by the proteomic analysis is related either to the regulation of angiogenesis/vasculogenesis or vascular homeostasis (Supporting Table S-6). A striking observation that further supports the endothelial commitment of the CD34$^+$ cells is the scarcity in CD34$^-$ cells of proteins that are implicated in vasculogenesis (Supporting Table S-6). This implies that the CD34$^+$ population is undoubtedly the one that eventually differentiates to endothelial cells. Among the few proteins with increased abundance in CD34$^-$ that might be involved in vasculogenesis were CREG1[74] LUM[75] and the LRP1[76] (Fig. 4E, Supporting Table S-7). Interestingly, CREG1 and LUM are secreted and LRP1 is cleaved[77] generating a soluble fragment. Therefore, even the CD34$^-$ cells might contribute to the vasculogenic differentaition of the CD34$^+$ cells via a potential cross-talk between these cell types mediated by diffusible molecules. Indeed, LRP1 acts as an endocytic receptor to Bmper and co-receptor of Bmp4 to mediate the endocytosis of the Mmper/Bmp4 signaling complex being a critical regulator of vascular development.[76] The newly identified proteins that are co-synthesized together with CD34 and the other known markers of endothelial commitment may contribute to a more accurate characterization of this vascular progenitor cell population in future studies.

In CD34$^+$ cells seven proteins that have been reported to be expressed in immature smooth muscle cells and play a role in vascular mural cell differentiation and function show increased abundance (PALD1, HDAC7, TLN1, VASP, CRIP2, LMNA, CORO1C) (Supporting Table S-6). This suggests that CD34$^+$ cells are bipotent: they can differentiate into either endothelial cells or vascular muscle cells (mural) depending on the growth conditions (Fig. 4B and C; Fig. 5, I and II), though concomitant differentiation to both lineages might be possible too. Indeed,

when we cultured CD34[+] cells with PDGFβ/TGFβ they exhibited a characteristic fibroblast-like morphology, they lost the expression of CD31 and up-regulated the mural marker SM22-alpha (Supporting Fig. S-2B). Thus, when fully *de novo* vessels are required, both VEGF and PDGF will have to be used to treat isolated CD34[+] progenitors from hiPSC patients, similarly to what was done to the hESC-derived CD34[+] progenitors of the present study. This treatment would drive the differentiation to new tight vessels that have both cell types in the correct proportion. In contrast, in diabetic retinopathy where the main abnormality is the lack of pericytes, then the CD34[+] progenitor cell (cultured on the correct scaffold) together with PDGF could be sufficient to remodel the damaged retina.[78]

Several cytoskeleton-related proteins are enriched in CD34[+] cells (Fig. 4D, Supporting Table S-6). Specifically, the ERM proteins (ezrin-EZR, radixin-RDX, moesin-MSN) consist of actin-binding linkers connecting F-actin and the plasma membrane, directly or indirectly.[79, 80] ERM proteins RDX and MSN and the kinase STK10 were over-synthesized in CD34[+] cells (Fig. 4D) as were the ERM interactors surface receptor CD44, the scaffold protein NHERF2 and the adaptor molecules CRK, VASP, TLN and PXN. Specific kinases, such as LOK/STK10, phosphorylate and activate the ERM proteins[79, 80] upon binding to plasma membrane PtsIns(4,5)P$_2$[81], or alternatively, upon binding to PDZ-domain proteins, such as NHERF2[82] (Fig. 5, right panel). Once activated, ERM proteins can bind one of several transmembrane receptors such as CD44[83] and can activate downstream kinases. The activated ERM proteins together with the other over-synthesized proteins regulate migration of vascular cells at many levels such as the cell-ECM contacts via integrins (PXN and TLN), association to CRK to promote migration[84] and/or the formation of tight bundles of synchronously polymerizing actin filaments by the leading edge filopodia with the participation of VASP proteins, which are also over-synthesised[85] (Fig. 5, right panel). ERM proteins, such as Ezrin, localize to the apical surfaces of many cells and are essential for establishing apical identity. Moreover, they are able to orient

the mitotic spindle to guide an asymmetric division that distinguishes the outer trophectoderm cells from unpolarized inner cell mass (ICM) cells.[86] MSN, which shows increased abundance in CD34$^+$ cells is involved in lumen formation in the newly formed vessels by delivering vesicles to the cadherin-based cell junctions to convert them to apical surfaces thereby hollowing the cord. This mechanism was validated in animal models both in mouse aorta and in zebrafish intersegmental vessels[87].

In the CD34$^-$ cells we have identified 5 proteins with increased abundance (CD140B, PALLD, CREG1, LRP1, DAG1; Fig. 4E) that are synthesized in vascular smooth muscle cells (VSMCs) or play a role in their differentiation and the proper investment of both large and small vessels with mural cells.[74, 77, 88-94] Importantly, hiPSCs-derived contractile and functional SMCs were CD34^{-73} indicating that both types of progenitor stem cell (CD34$^+$ and CD34$^-$) have the machinery to differentiate SMCs/PCs. This is in agreement with the fact that during development, VSMCs arise from multiple independent origins or different subsets of mesoderm-committed cells.[95]

The presence of SMC/PC progenitors that are CD34$^-$ is raising the question about their possible contribution in vessel morphogenesis considering that this stem cell population is devoid of commitment to an endothelial cell lineage. In this sense it is not a vascular progenitor cell proper as it can differentiate only into the mural component of it. Certainly, contribution to vessel formation impacted by the CD34$^+$ progenitors is an obvious explanation. However, the exact nature of such contribution awaits further future characterization. The isolated CD34$^-$ cells are a very interesting and important, from the clinical translation point of view, group of progenitor populations. The pattern of proteins with up-regulated synthesis in CD34$^-$ cells is consistent with differentiation of mesoderm to cell populations of the mesenchyme-direction that consist of skeletal myogenic /osteogenic/chondrogenic/adipogenic lineages (Supporting Table S-7)(Fig. 5,III) suggesting that in CD34$^-$ cells there is either a common progenitor that

differentiates to all these lineages or one progenitor for each lineage or combinations thereof.

Many of the 51 proteins showing increased abundance in this cell population, including 6

membrane proteins (ATP2B1, CRB2, LRP2, MPZL1, NCAM1, PTK7, Supporting Table S-4),

may prove useful as individual markers of the various cell-subtypes and as surface tags for

sorting.

Conclusion

In conclusion, by modifying an established method, we differentiated hESCs to distinct CD34$^+$ and CD34$^-$ cells that could be segregated from each other. CD34$^+$ cells were induced to differentiate towards the endothelial lineage by VEGF, as evidenced by validation using known markers and functional characterization (Fig. 5I), whereas PDGF/TGFβ1 induced the expression of mural cell (SMCs/PCs) markers (Fig. 5II). Thus, CD34$^+$ cells probably represent a bipotent progenitor population that can differentiate either to ECs or SMCs/PCs. Comparative proteomic supported this conclusion as most of the specifically over-synthesized proteins in hESC-derived CD34$^+$ cells were involved in the regulation of vasculogenesis/angiogenesis, including also proteins that have been reported to be expressed in immature differentiating SMCs. On the contrary, CD34$^-$ cells are strikingly devoid of proteins that have any direct regulatory role in vasculogenesis/angiogenesis, but do over-synthesize a few proteins that drive mural cell differentiation (Fig. 5II). However, the pattern of over-synthesized proteins in CD34$^-$ cells is consistent with differentiation to the mesenchyme-direction that consist of skeletal myogenic /osteogenic/chondrogenic/adipogenic lineages (Fig. 5III). The identified proteins due to the rigorous criteria used may serve as markers for further characterization of subpopulations of these cells in future studies.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

KK and KT performed the proteomic experiments and analysis and drafted the manuscript; EB isolated, cultured induced the differentiation of human embryonic stem cells isolated the vascular progenitor cells, performed the characterization validation and functional characterization and drafted sections of the manuscript; SZ participated in the isolation and the characterization of vascular progenitor cells; ESB and MZ contributed to the functional annotation analysis; SC contributed and guided the statistical analysis; SCh and CM participated in guiding the cell biology experiments and the interpretation of the results; TF conceived the study, guided the cell biology experiments, interpreted results and edited the manuscript, AE conceived the study, guided the proteomics experiments and analyses, interpreted results, drafted and did the final editing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

References

1.   Langer, R.; Vacanti, J. P., Tissue Engineering. *Science* 1993, 260, 920-926.

2.   Hughes, C. S.; Nuhn, A. A.; Postovit, L. M.; Lajoie, G. A., Proteomics of human embryonic stem cells. *Proteomics* 2011, 11, 675-90.

3.   Van Hoof, D.; Heck, A. J.; Krijgsveld, J.; Mummery, C. L., Proteomics and human embryonic stem cells. *Stem cell research* 2008, 1, 169-82.

4.   Takahashi, K.; Yamanaka, S., Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006, 126, 663-676.

5.   Novosel, E. C.; Kleinhans, C.; Kluger, P. J., Vascularization is the key challenge in tissue engineering. *Adv Drug Deliver Rev* 2011, 63, 300-311.

6.   Phelps, E. A.; Garcia, A. J., Engineering more than a cell: vascularization strategies in tissue engineering. *Curr Opin Biotech* 2010, 21, 704-709.

7.   Jain, R. K., Normalization of tumor vasculature: An emerging concept in antiangiogenic therapy. *Science* 2005, 307, 58-62.

8.   Carmeliet, P.; Jain, R. K., Angiogenesis in cancer and other diseases. *Nature* 2000, 407, 249-257.

9.   Lu, L. C.; Arbit, H. M.; Herrick, J. L.; Segovis, S. G.; Maran, A.; Yaszemski, M. J., Tissue Engineered Constructs: Perspectives on Clinical Translation. *Ann Biomed Eng* 2015, 43, 796-804.

10.   Tsigkou, O.; Pomerantseva, I.; Spencer, J. A.; Redondo, P. A.; Hart, A. R.; O'Doherty, E.; Lin, Y. F.; Friedrich, C. C.; Daheron, L.; Lin, C. P.; Sundback, C. A.; Vacanti, J. P.; Neville, C., Engineered vascularized bone grafts. *P Natl Acad Sci USA* 2010, 107, 3311-3316.

11.   Zhou, Y.; Yang, F.; Chen, T.; Wu, Y.; Yang, M.; Zhu, J.; Zhang, L., An updated view on the differentiation of stem cells into endothelial cells. *Science China. Life sciences* 2014, 57, 763-73.

12.   Orlova, V. V.; Drabsch, Y.; Freund, C.; Petrus-Reurer, S.; van den Hil, F. E.; Muenthaisong, S.; Dijke, P. T.; Mummery, C. L., Functionality of endothelial cells and pericytes from human pluripotent stem cells demonstrated in cultured vascular plexus and zebrafish xenografts. *Arteriosclerosis, thrombosis, and vascular biology* 2014, 34, 177-86.

13.   Rufaihah, A. J.; Huang, N. F.; Kim, J.; Herold, J.; Volz, K. S.; Park, T. S.; Lee, J. C.; Zambidis, E. T.; Reijo-Pera, R.; Cooke, J. P., Human induced pluripotent stem cell-derived endothelial cells exhibit functional heterogeneity. *American journal of translational research* 2013, 5, 21-35.

14.   Park, T. S.; Bhutto, I.; Zimmerlin, L.; Huo, J. S.; Nagaria, P.; Miller, D.; Rufaihah, A. J.; Talbot, C.; Aguilar, J.; Grebe, R.; Merges, C.; Reijo-Pera, R.; Feldman, R. A.; Rassool, F.; Cooke, J.; Lutty, G.; Zambidis, E. T., Vascular progenitors from cord blood-derived induced pluripotent stem cells possess augmented capacity for regenerating ischemic retinal vasculature. *Circulation* 2014, 129, 359-72.

15.   Patsch, C.; Challet-Meylan, L.; Thoma, E. C.; Urich, E.; Heckel, T.; O'Sullivan, J. F.; Grainger, S. J.; Kapp, F. G.; Sun, L.; Christensen, K.; Xia, Y.; Florido, M. H.; He, W.; Pan, W.; Prummer, M.; Warren, C. R.; Jakob-Roetne, R.; Certa, U.; Jagasia, R.; Freskgard, P. O.; Adatto, I.; Kling, D.; Huang, P.; Zon, L. I.; Chaikof, E. L.; Gerszten, R. E.; Graf, M.; Iacone, R.; Cowan, C. A., Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. *Nat Cell Biol* 2015, 17, 994-1003.

16.   Lee, J. B.; Werbowetski-Ogilvie, T. E.; Lee, J. H.; McIntyre, B. A.; Schnerch, A.; Hong, S. H.; Park, I. H.; Daley, G. Q.; Bernstein, I. D.; Bhatia, M., Notch-HES1 signaling axis controls hemato-endothelial fate decisions of human embryonic and induced pluripotent stem cells. *Blood* 2013, 122, 1162-73.

17.   Costa, M.; Sourris, K.; Lim, S. M.; Yu, Q. C.; Hirst, C. E.; Parkington, H. C.; Jokubaitis, V. J.; Dear, A. E.; Liu, H. B.; Micallef, S. J.; Koutsis, K.; Elefanty, A. G.; Stanley, E. G., Derivation of endothelial cells from human embryonic stem cells in fully defined medium enables identification of lysophosphatidic acid and platelet activating factor as regulators of eNOS localization. *Stem cell research* 2013, 10, 103-17.

18.   White, M. P.; Rufaihah, A. J.; Liu, L.; Ghebremariam, Y. T.; Ivey, K. N.; Cooke, J. P.; Srivastava, D., Limited gene expression variation in human embryonic stem cell and induced pluripotent stem cell-derived endothelial cells. *Stem Cells* 2013, 31, 92-103.

19.   Li, Z.; Hu, S.; Ghosh, Z.; Han, Z.; Wu, J. C., Functional characterization and expression profiling of human induced pluripotent stem cell- and embryonic stem cell-derived endothelial cells. *Stem Cells Dev* 2011, 20, 1701-10.

20.   Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M., Global quantification of mammalian gene expression control. *Nature* 2011, 473, 337-42.

21.   Larance, M.; Lamond, A. I., Multidimensional proteomics for cell biology. *Nature reviews. Molecular cell biology* 2015, 16, 269-80.

22.   Phanstiel, D. H.; Brumbaugh, J.; Wenger, C. D.; Tian, S.; Probasco, M. D.; Bailey, D. J.; Swaney, D. L.; Tervo, M. A.; Bolin, J. M.; Ruotti, V.; Stewart, R.; Thomson, J. A.; Coon, J. J., Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nature methods* 2011, 8, 821-7.

23.   Bohman, S.; Matsumoto, T.; Suh, K.; Dimberg, A.; Jakobsson, L.; Yuspa, S.; Welsh, L. C., Proteomic analysis of vascular endothelial growth factor-induced endothelial cell differentiation reveals a role for chloride intracellular channel 4 (CLIC4) in tubular morphogenesis. *J Biol Chem* 2005, 280, 42397-42404.

24.   Trakarnsanga, K.; Wilson, M. C.; Griffiths, R. E.; Toye, A. M.; Carpenter, L.; Heesom, K. J.; Parsons, S. F.; Anstee, D. J.; Frayne, J., Qualitative and Quantitative Comparison of the Proteome of Erythroid Cells Differentiated from Human iPSCs and Adult Erythroid Cells by Multiplex TMT Labelling and NanoLC-MS/MS. *Plos One* 2014, 9.

25.   Liu, F.; Lu, J.; Fang, H. H.; Wang, Z. Q.; Cui, S. J.; Zhang, G. A.; Chi, M.; Zhang, X.; Yang, P. Y.; Chen, Z.; Han, Z. G., Insights into human CD34(+) hematopoietic stem/progenitor cells through a systematically proteomic survey coupled with transcriptome. *Proteomics* 2006, 6, 2673-2692.

26.   Prokopi, M.; Pula, G.; Mayr, U.; Devue, C.; Gallagher, J.; Xiao, Q.; Boulanger, C. M.; Westwood, N.; Urbich, C.; Willeit, J.; Steiner, M.; Breuss, J.; Xu, Q. B.; Kiechl, S.; Mayr, M., Proteomic analysis reveals presence of platelet microparticles in endothelial progenitor cell cultures. *Blood* 2009, 114, 723-732.

27.   Bahramsoltani, M.; Harms, T.; Drewes, B.; Plendl, J., Searching for markers to identify angiogenic endothelial cells: A proteomic approach. *Clin Hemorheol Micro* 2013, 55, 255-269.

28.   Tan, J. Y.; Sriram, G.; Rufaihah, A. J.; Neoh, K. G.; Cao, T., Efficient Derivation of Lateral Plate and Paraxial Mesoderm Subtypes from Human Embryonic Stem Cells Through GSKi-Mediated Differentiation. *Stem Cells Dev* 2013, 22, 1893-1906.

29.   Suwinska, A.; Ciemerych, M. A., Factors Regulating Pluripotency and Differentiation in Early Mammalian Embryos and Embryo-Derived Stem Cells. *Vitam Horm* 2011, 87, 1-37.

30.   Ng, E. S.; Davis, R.; Stanley, E. G.; Elefanty, A. G., A protocol describing the use of a recombinant protein-based, animal product-free medium (APEL) for human embryonic stem cell differentiation as spin embryoid bodies. *Nat Protoc* 2008, 3, 768-776.

31.   Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 2008, 26, 1367-72.

32.   Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M., Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research* 2011, 10, 1794-805.

33.   Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics : MCP* 2014, 13, 2513-26.

34.   Vizcaino, J. A.; Csordas, A.; Del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H., 2016 update of the PRIDE database and its related tools. *Nucleic acids research* 2016, 44, D447-56.

35.   Wold, S.; Sjostrom, M.; Eriksson, L., PLS-regression: a basic tool of chemometrics. *Chemometr Intell Lab* 2001, 58, 109-130.

36.   Carpentier, S. C., Multiple Testing and Pattern Recognition in 2-DE Proteomics. *Methods in molecular biology (Clifton, N.J.)* 2016, 1384, 215-35.

37.   Mehmood, T.; Liland, K. H.; Snipen, L.; Saebo, S., A review of variable selection methods in Partial Least Squares Regression. *Chemometr Intell Lab* 2012, 118, 62-69.

38.   Sun, X.; Cheng, L.; Duan, H.; Lin, G.; Lu, G., Characterization and comparison of embryonic stem cell-derived KDR+ cells with endothelial cells. *Microvasc Res* 2012, 84, 149-54.

39.   Marcelo, K. L.; Goldie, L. C.; Hirschi, K. K., Regulation of endothelial cell differentiation and specification. *Circ Res* 2013, 112, 1272-87.

40.   Ditadi, A.; Sturgeon, C. M.; Tober, J.; Awong, G.; Kennedy, M.; Yzaguirre, A. D.; Azzola, L.; Ng, E. S.; Stanley, E. G.; French, D. L.; Cheng, X.; Gadue, P.; Speck, N. A.; Elefanty, A. G.; Keller, G., Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nat Cell Biol* 2015, 17, 580-91.

41.   Lee, D.; Park, C.; Lee, H.; Lugus, J. J.; Kim, S. H.; Arentson, E.; Chung, Y. S.; Gomez, G.; Kyba, M.; Lin, S.; Janknecht, R.; Lim, D. S.; Choi, K., ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. *Cell Stem Cell* 2008, 2, 497-507.

42.   Sumanas, S.; Lin, S., Ets1-related protein is a key regulator of vasculogenesis in zebrafish. *PLoS Biol* 2006, 4, e10.

43.   Salanga, M. C.; Meadows, S. M.; Myers, C. T.; Krieg, P. A., ETS family protein ETV2 is required for initiation of the endothelial lineage but not the hematopoietic lineage in the Xenopus embryo. *Dev Dyn* 2010, 239, 1178-87.

44.   Wong, K. S.; Proulx, K.; Rost, M. S.; Sumanas, S., Identification of vasculature-specific genes by microarray analysis of Etsrp/Etv2 overexpressing zebrafish embryos. *Dev Dyn* 2009, 238, 1836-50.

45.   Gomez, G. A.; Veldman, M. B.; Zhao, Y.; Burgess, S.; Lin, S., Discovery and characterization of novel vascular and hematopoietic genes downstream of etsrp in zebrafish. *Plos One* 2009, 4, e4994.

46.   Armulik, A.; Abramsson, A.; Betsholtz, C., Endothelial/pericyte interactions. *Circ Res* 2005, 97, 512-23.

47.   Li, L.; Miano, J. M.; Cserjesi, P.; Olson, E. N., SM22 alpha, a marker of adult smooth muscle, is expressed in multiple myogenic lineages during embryogenesis. *Circ Res* 1996, 78, 188-95.

48.   Beck, M.; Schmidt, A.; Malmstroem, J.; Claassen, M.; Ori, A.; Szymborska, A.; Herzog, F.; Rinner, O.; Ellenberg, J.; Aebersold, R., The quantitative proteome of a human cell line. *Molecular systems biology* 2011, 7, 549.

49.   Geiger, T.; Wehner, A.; Schaab, C.; Cox, J.; Mann, M., Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Molecular & cellular proteomics : MCP* 2012, 11, M111.014050.

50.   Wilhelm, M.; Schlegl, J.; Hahne, H.; Moghaddas Gholami, A.; Lieberenz, M.; Savitski, M. M.; Ziegler, E.; Butzmann, L.; Gessulat, S.; Marx, H.; Mathieson, T.; Lemeer, S.; Schnatbaum, K.; Reimer, U.; Wenschuh, H.; Mollenhauer, M.; Slotta-Huspenina, J.; Boese, J. H.; Bantscheff, M.; Gerstmair, A.; Faerber, F.; Kuster, B., Mass-spectrometry-based draft of the human proteome. *Nature* 2014, 509, 582-7.

51.   Boulesteix, A. L.; Strimmer, K., Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Briefings in bioinformatics* 2007, 8, 32-44.

52.   Rodda, D. J.; Chew, J. L.; Lim, L. H.; Loh, Y. H.; Wang, B.; Ng, H. H.; Robson, P., Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* 2005, 280, 24731-7.

53.   Gaspar-Maia, A.; Alajem, A.; Polesso, F.; Sridharan, R.; Mason, M. J.; Heidersbach, A.; Ramalho-Santos, J.; McManus, M. T.; Plath, K.; Meshorer, E.; Ramalho-Santos, M., Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* 2009, 460, 863-8.

54.   Hou, Z.; Meyer, S.; Propson, N. E.; Nie, J.; Jiang, P.; Stewart, R.; Thomson, J. A., Characterization and target identification of a DNA aptamer that labels pluripotent stem cells. *Cell research* 2015, 25, 390-3.

55.   Birdsey, G. M.; Shah, A. V.; Dufton, N.; Reynolds, L. E.; Osuna Almagro, L.; Yang, Y.; Aspalter, I. M.; Khan, S. T.; Mason, J. C.; Dejana, E.; Gottgens, B.; Hodivala-Dilke, K.; Gerhardt, H.; Adams, R. H.; Randi, A. M., The endothelial transcription factor ERG promotes vascular stability and growth through Wnt/beta-catenin signaling. *Developmental cell* 2015, 32, 82-96.

56.   Nassiri, F.; Cusimano, M. D.; Scheithauer, B. W.; Rotondo, F.; Fazio, A.; Yousef, G. M.; Syro, L. V.; Kovacs, K.; Lloyd, R. V., Endoglin (CD105): a review of its role in angiogenesis and tumor diagnosis, progression and therapy. *Anticancer Res* 2011, 31, 2283-90.

57.   Crisan, M.; Chen, C. W.; Corselli, M.; Andriolo, G.; Lazzari, L.; Peault, B., Perivascular multipotent progenitor cells in human organs. *Ann N Y Acad Sci* 2009, 1176, 118-23.

58.   Carbon, S.; Ireland, A.; Mungall, C. J.; Shu, S.; Marshall, B.; Lewis, S.; Hub, A.; Grp, W. P. W., AmiGO: online access to ontology and annotation data. *Bioinformatics* 2009, 25, 288-289.

59.   Palmer, R. M.; Ashton, D. S.; Moncada, S., Vascular endothelial cells synthesize nitric oxide from L-arginine. *Nature* 1988, 333, 664-6.

60.   Huang, N. F.; Fleissner, F.; Sun, J.; Cooke, J. P., Role of Nitric Oxide Signaling in Endothelial Differentiation of Embryonic Stem Cells. *Stem Cells Dev* 2010, 19, 1617-1625.

61.   Zhang, B.; Kirov, S.; Snoddy, J., WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic acids research* 2005, 33, W741-8.

62.   Wang, J.; Duncan, D.; Shi, Z.; Zhang, B., WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic acids research* 2013, 41, W77-83.

63.   Wallgard, E.; Nitzsche, A.; Larsson, J.; Guo, X.; Dieterich, L. C.; Dimberg, A.; Olofsson, T.; Ponten, F. C.; Makinen, T.; Kalen, M.; Hellstrom, M., Paladin (X99384) is expressed in the vasculature and shifts from endothelial to vascular smooth muscle cells during mouse development. *Dev Dyn* 2012, 241, 770-86.

64.   Nillesen, S. T.; Geutjes, P. J.; Wismans, R.; Schalkwijk, J.; Daamen, W. F.; van Kuppevelt, T. H., Increased angiogenesis and blood vessel maturation in acellular collagen-heparin scaffolds containing both FGF2 and VEGF. *Biomaterials* 2007, 28, 1123-31.

65.   Li, Z.; Wilson, K. D.; Smith, B.; Kraft, D. L.; Jia, F.; Huang, M.; Xie, X.; Robbins, R. C.; Gambhir, S. S.; Weissman, I. L.; Wu, J. C., Functional and transcriptional characterization of human embryonic stem cell-derived endothelial cells for treatment of myocardial infarction. *Plos One* 2009, 4, e8443.

66.   Rufaihah, A. J.; Huang, N. F.; Jame, S.; Lee, J. C.; Nguyen, H. N.; Byers, B.; De, A.; Okogbaa, J.; Rollins, M.; Reijo-Pera, R.; Gambhir, S. S.; Cooke, J. P., Endothelial cells derived

32

from human iPSCS increase capillary density and improve perfusion in a mouse model of peripheral arterial disease. *Arteriosclerosis, thrombosis, and vascular biology* 2011, 31, e72-9.

67.   Samuel, R.; Daheron, L.; Liao, S.; Vardam, T.; Kamoun, W. S.; Batista, A.; Buecker, C.; Schafer, R.; Han, X.; Au, P.; Scadden, D. T.; Duda, D. G.; Fukumura, D.; Jain, R. K., Generation of functionally competent and durable engineered blood vessels from human induced pluripotent stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 2013, 110, 12774-9.

68.   Koike, N.; Fukumura, D.; Gralla, O.; Au, P.; Schechner, J. S.; Jain, R. K., Tissue engineering: creation of long-lasting blood vessels. *Nature* 2004, 428, 138-9.

69.   Au, P.; Daheron, L. M.; Duda, D. G.; Cohen, K. S.; Tyrrell, J. A.; Lanning, R. M.; Fukumura, D.; Scadden, D. T.; Jain, R. K., Differential in vivo potential of endothelial progenitor cells from human umbilical cord blood and adult peripheral blood to form functional long-lasting vessels. *Blood* 2008, 111, 1302-5.

70.   Wang, Z. Z.; Au, P.; Chen, T.; Shao, Y.; Daheron, L. M.; Bai, H.; Arzigian, M.; Fukumura, D.; Jain, R. K.; Scadden, D. T., Endothelial cells derived from human embryonic stem cells form durable blood vessels in vivo. *Nature biotechnology* 2007, 25, 317-8.

71.   Au, P.; Tam, J.; Fukumura, D.; Jain, R. K., Bone marrow-derived mesenchymal stem cells facilitate engineering of long-lasting functional vasculature. *Blood* 2008, 111, 4551-8.

72.   Carmeliet, P.; Jain, R. K., Molecular mechanisms and clinical applications of angiogenesis. *Nature* 2011, 473, 298-307.

73.   Bajpai, V. K.; Mistriotis, P.; Loh, Y. H.; Daley, G. Q.; Andreadis, S. T., Functional vascular smooth muscle cells derived from human induced pluripotent stem cells via mesenchymal stem cell intermediates. *Cardiovascular research* 2012, 96, 391-400.

74.   Tian, X.; Zhang, N.; Yan, C.; Nelsen, J.; Liu, S.; Kang, J.; Zhang, J.; Peng, C.; Tao, J.; Sun, M.; Li, Y.; Li, S.; Wang, H.; Han, Y., CREG promotes vasculogenesis by activation of VEGF/PI3K/Akt pathway. *Front Biosci (Landmark Ed)* 2014, 19, 1215-26.

75.   Albig, A. R.; Roy, T. G.; Becenti, D. J.; Schiemann, W. P., Transcriptome analysis of endothelial cell gene expression induced by growth on matrigel matrices: identification and characterization of MAGP-2 and lumican as novel regulators of angiogenesis. *Angiogenesis* 2007, 10, 197-216.

76.   Pi, X.; Schmitt, C. E.; Xie, L.; Portbury, A. L.; Wu, Y.; Lockyer, P.; Dyer, L. A.; Moser, M.; Bu, G.; Flynn, E. J., 3rd; Jin, S. W.; Patterson, C., LRP1-dependent endocytic mechanism governs the signaling output of the bmp system in endothelial cells and in angiogenesis. *Circ Res* 2012, 111, 564-74.

77.   Boucher, P.; Herz, J., Signaling through LRP1: Protection from atherosclerosis and beyond. *Biochem Pharmacol* 2011, 81, 1-5.

78.   Beltramo, E.; Porta, M., Pericyte loss in diabetic retinopathy: mechanisms and consequences. *Current medicinal chemistry* 2013, 20, 3218-25.

79.   Wang, Y.; Kaiser, M. S.; Larson, J. D.; Nasevicius, A.; Clark, K. J.; Wadman, S. A.; Roberg-Perez, S. E.; Ekker, S. C.; Hackett, P. B.; McGrail, M.; Essner, J. J., Moesin1 and Ve-cadherin are required in endothelial cells during in vivo tubulogenesis. *Development* 2010, 137, 3119-28.

80.   Belkina, N. V.; Liu, Y.; Hao, J. J.; Karasuyama, H.; Shaw, S., LOK is a major ERM kinase in resting lymphocytes and regulates cytoskeletal rearrangement through ERM phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* 2009, 106, 4707-12.

81.   Fehon, R. G.; McClatchey, A. I.; Bretscher, A., Organizing the cell cortex: the role of ERM proteins. *Nature reviews. Molecular cell biology* 2010, 11, 276-87.

Tsolis et al   *Endothelial transition proteomics*   33

82.   Viswanatha, R.; Ohouo, P. Y.; Smolka, M. B.; Bretscher, A., Local phosphocycling mediated by LOK/SLK restricts ezrin function to the apical aspect of epithelial cells. *J Cell Biol* 2012, 199, 969-84.

83.   Yonemura, S.; Hirao, M.; Doi, Y.; Takahashi, N.; Kondo, T.; Tsukita, S.; Tsukita, S., Ezrin/radixin/moesin (ERM) proteins bind to a positively charged amino acid cluster in the juxta-membrane cytoplasmic domain of CD44, CD43, and ICAM-2. *J Cell Biol* 1998, 140, 885-95.

84.   Boratko, A.; Csortos, C., NHERF2 is crucial in ERM phosphorylation in pulmonary endothelial cells. *Cell Commun Signal* 2013, 11, 99.

85.   Tsuda, M.; Makino, Y.; Iwahara, T.; Nishihara, H.; Sawa, H.; Nagashima, K.; Hanafusa, H.; Tanaka, S., Crk associates with ERM proteins and promotes cell motility toward hyaluronic acid. *J Biol Chem* 2004, 279, 46843-50.

86.   Barreiro, O.; Yanez-Mo, M.; Serrador, J. M.; Montoya, M. C.; Vicente-Manzanares, M.; Tejedor, R.; Furthmayr, H.; Sanchez-Madrid, F., Dynamic interaction of VCAM-1 and ICAM-1 with moesin and ezrin in a novel endothelial docking structure for adherent leukocytes. *J Cell Biol* 2002, 157, 1233-45.

87.   Adyshev, D. M.; Dudek, S. M.; Moldobaeva, N.; Kim, K. M.; Ma, S. F.; Kasa, A.; Garcia, J. G.; Verin, A. D., Ezrin/radixin/moesin proteins differentially regulate endothelial hyperpermeability after thrombin. *Am J Physiol Lung Cell Mol Physiol* 2013, 305, L240-55.

88.   Nguyen, N. U.; Wang, H. V., Dual roles of palladin protein in in vitro myogenesis: inhibition of early induction but promotion of myotube maturation. *Plos One* 2015, 10, e0124762.

89.   Jin, L.; Gan, Q.; Zieba, B. J.; Goicoechea, S. M.; Owens, G. K.; Otey, C. A.; Somlyo, A. V., The actin associated protein palladin is important for the early smooth muscle cell differentiation. *Plos One* 2010, 5, e12823.

90.   Nakajima, C.; Haffner, P.; Goerke, S. M.; Zurhove, K.; Adelmann, G.; Frotscher, M.; Herz, J.; Bock, H. H.; May, P., The lipoprotein receptor LRP1 modulates sphingosine-1-phosphate signaling and is essential for vascular development. *Development* 2014, 141, 4513-25.

91.   Moiseeva, E. P., Adhesion receptors of vascular smooth muscle cells and their functions. *Cardiovascular research* 2001, 52, 372-86.

92.   Cohn, R. D.; Henry, M. D.; Michele, D. E.; Barresi, R.; Saito, F.; Moore, S. A.; Flanagan, J. D.; Skwarchuk, M. W.; Robbins, M. E.; Mendell, J. R.; Williamson, R. A.; Campbell, K. P., Disruption of DAG1 in differentiated skeletal muscle reveals a role for dystroglycan in muscle regeneration. *Cell* 2002, 110, 639-48.

93.   Cote, P. D.; Moukhles, H.; Lindenbaum, M.; Carbonetto, S., Chimaeric mice deficient in dystroglycans develop muscular dystrophy and have disrupted myoneural synapses. *Nat Genet* 1999, 23, 338-42.

94.   Montanaro, F.; Lindenbaum, M.; Carbonetto, S., alpha-Dystroglycan is a laminin receptor involved in extracellular matrix assembly on myotubes and muscle cell viability. *J Cell Biol* 1999, 145, 1325-40.

95.   Cheung, C.; Bernardo, A. S.; Trotter, M. W.; Pedersen, R. A.; Sinha, S., Generation of human vascular smooth muscle subtypes provides insight into embryological origin-dependent disease susceptibility. *Nature biotechnology* 2012, 30, 165-73.

Illustrations and figures

Figure 1. Cell differentiation protocol clone characterization.

A) hESCs were differentiated in CD34[+] cells, under feeder-free conditions in matrigel-coated plates, using a modified protocol of Tan et al. 2013.[28] B) FACS analysis of CD34 and CD31 expression on Day 0 (left column) and on Day 4 (right column) of the differentiation process, before CD34 magnetic selection. The histograms of the samples stained with the isotypic IgGs are shown in light gray, whereas the samples stained with fluorochrome-conjugated antibodies are overlaid in black. Each histogram is a representative of at least ten separate experiments. C) FACS analysis of CD34 and CD31 expression in CD34[+] and CD34[-] cell fraction after CD34 magnetic selection. Representative data from at least ten independed experiments are presented as dot plots (FITC, fluorescein isothiocyanate; PE, phycoerythrin). D) Matrigel tube formation assay: Representative photographs of CD34[+] and CD34[-] selected cells after seeding for 18 hours on Matrigel. Capillary-like structures were imaged on Zeiss Axiovert 100 using a 5 X objective. E) Cell differentiation was also validated by examining several additional markers. The theoretical phenotype is in agreement with our experimental data derived from W (Western Blot), F (FACS), I (Immunofluoresence microcopy) and mass spectrometry. Additional markers are shown in Supporting Figures S-1 and S-2.

Figure 2. Proteomics analysis workflow

In the present study we performed a proteomic characterization of pluripotent hESCs and the two differentiated cell populations of CD34[+] VPCs and CD34[-] cells using high resolution mass spectrometry combined with label-free quantification and multivariate analysis. 6 biological repeats for each cell type were analyzed resulting in the identification of ~4.500

proteins. A) Plot of the scores from the PLS model for the two main principal components. Each

dot represents one biological sample. Three distinct clusters are identified, containing the 6

biological repeats of each cell type. The distance between hESCs and CD34$^{+/-}$ cells is greater

than the distance between CD34^{+} and CD34^{-} suggesting a bigger variation of the starting

population with the two differentiated cell types. B) Plot of the loadings (identified proteins)

from the PLS model for the two main PCs. For the selection of differential proteins we used

VIP method to select proteins with the highest variation between the different cell types, and

checked for statistical significance using the non-parametric test Kruskal-Wallis. 236 proteins

show differential synthesis between hESCs, CD34^{+} and CD34^{-} cells. A part of the differentially

synthesized proteins is presented in the plot (for complete list see Supporting Table S-3). C)

Distribution of differential synthesized proteins between the study groups. D) Differential

proteins were then used for the biological characterization of cells using GO-slim, pathway

enrichment and manual curation.

Figure 3. GO enrichment analysis

Differentially synthesized proteins were subjected to GO-slim analysis using AmiGO. hESCs show higher number of proteins related with chromosome or methylation events. CD34[+] cells synthesized more proteins associated with cytoskeleton, adhesion, locomotion or are related to membrane-bound vesicles and vesicle mediated transport. The CD34[-] population comprise an intermediate phenotype between hESCs and CD34[+] cells, but shows significantly less proteins compared to the CD34[+] cells that are related with adhesion/cytoskeleton and membrane vesicle transport.

Figure 4. Selected proteins of functional interest.

A) Proteomics results are validated using additional methods. The markers OCT4, ALPL, KDR, CD144, CD201, ERG, CD44 and CD140B, show statistically significant differential abundance based on the proteomics results, similar with the immunological methods (Fig. 1, Supporting Fig. S-1). B) Representative set of proteins abundant in CD34[+] cells that are related with vasculogenesis (see Supporting Table S-5 for additional proteins). C) Abundant proteins in CD34[+] cells that participate in muscle cell differentiation. D) ERM proteins and proteins associated with ERM family show higher quantitative value in CD34[+] cells, suggesting a role of ERM protein family in physiological cell function. E) CD34[-] cells, consist of a mixed population. Three secreted proteins (CREG1, LUM, LRP1) are more abundant in CD34[-] cells suggesting a possible paracrine effect of these cells to CD34[+]. In addition, proteins related to muscle cell differentiation, bone/cartilage formation and adipogenic differentiation are also present in CD34[-] population. Representative proteins are illustrated in this panel (see Supporting Table S-6 for the complete list).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

38

Figure 5. Model of transition of hESCs into differentiated cells.

hESCs differentiate into mesoderm intermediates and then into CD34$^+$ VPCs and CD34$^-$ cells, under feeder-free conditions, using defined media. CD34$^+$ synthesize known vascular cell markers, as well as proteins related to vasculogenesis and smooth muscle differentiation, suggesting a bipotent phenotype. Indeed, CD34$^+$ can differentiate further to both directions, depending on the growth conditions (Supporting Fig. S-2). CD34$^-$ cells consist of a mixed population over-synthesizing secreted factors that can promote angiogenesis, suggesting a paracrine effect on CD34$^+$ cells and also synthesizing proteins that can lead to differentiation to other lineages (muscle, bone, cartilage).

Tables and captions
Table 1. Enriched pathways.

Differential proteins were subjected in pathway enrichment analysis against Wikipathways database, using WebGestalt web tool. Enriched pathways can be grouped in three main categories, i) adhesion and mechanotransduction, ii) signaling pathways and iii) regulatory and metabolic pathways.

Pathway name	# proteins	Protein gene names	Adjusted p-value
i) Adhesion - Mechanotransduction			
Integrin-mediated cell adhesion	8	VAV3 SRBS1 PAXI TLN1 ITA6 VASP CRK ITA9	7.71E-07
Regulation of Actin Cytoskeleton	7	PAXI GIT1 GELS CRK RADI MOES PGFRB	0.0001
Focal Adhesion	13	CO6A2 PAXI LAMC1 TLN1 VGFR2 ITA6 LAMA1 VASP CRK LAMB2 CO1A2 ITA9 PGFRB	4.56E-10
ii) Signaling pathways			
FAS pathway and Stress induction of HSP regulation	3	LMNA HSPB1 SPTAN1	0.008
Insulin Signaling	5	SORBS1 MYO1C KIF5B CRK EHD2	0.0062
AGE-RAGE pathway	3	ALPL MSN NOS3	0.0177
MAPK signaling pathway	4	CRK PPP5C HSPB1 PDGFRB	0.0226
Signaling of Hepatocyte Growth Factor Receptor	2	PXN CRK	0.0276
iii) Regulatory and metabolic pathways			
Glycolysis and Gluconeogenesis	4	PGK1 ENO3 HK2 HK1	0.0008
Urea cycle and metabolism of amino groups	3	OAT CKB GATM	0.0008
Angiogenesis	2	KDR NOS3	0.0177
miRNA regulation of DNA Damage Response	3	CCNB1 PML CDK6	0.0177
Glucuronidation	2	UGP2 HK1	0.0177
Prostaglandin Synthesis and Regulation	2	ANXA3 ANXA2	0.0225
One Carbon Metabolism	2	MTR DNMT3B	0.0276

Supporting Information components

Supporting Experimental Procedures:

Isolation of CD34$^+$ cells

Matrigel tube formation assay

Flow Cytometry

Immunofluorescence analysis

Western blot analysis

Quantitative RT-PCR (qRT-PCR)

Protein extraction from H1 hESCs, CD34$^+$ and CD34$^-$ cells

1D-SDS-PAGE and in-gel digestion

GO slim and pathway enrichment analysis

Supporting Figures

Supporting Figure S-1 – Phenotypic characterization of CD34+ cells

Supporting Figure S-2 – Differentiation potential of CD34+ cells.

Supporting Figure S-3 – Label-free quantification reproducibility and identification depth.

Supporting Figure S-4 – Visualization of differentially synthesized proteins.

Supporting Tables

Supporting Table S-1 – List of Antibodies used for immunophenotyping.

Supporting Table S-2 – qRT-PCR primers.

Supporting Table S-3 –Proteins identified by MS.

Supporting Table S-4 – Differentially synthesized proteins.

Supporting Table S-5 – GO-slim analysis.

Supporting Table S-6 – Known biological roles of proteins that are significantly over-synthesized in CD34$^+$ cells, with corresponding literature references.

Supporting Table S-7 – Known biological roles of proteins that are significantly over-synthesized in CD34- cells, with corresponding literature references.

Supporting Files

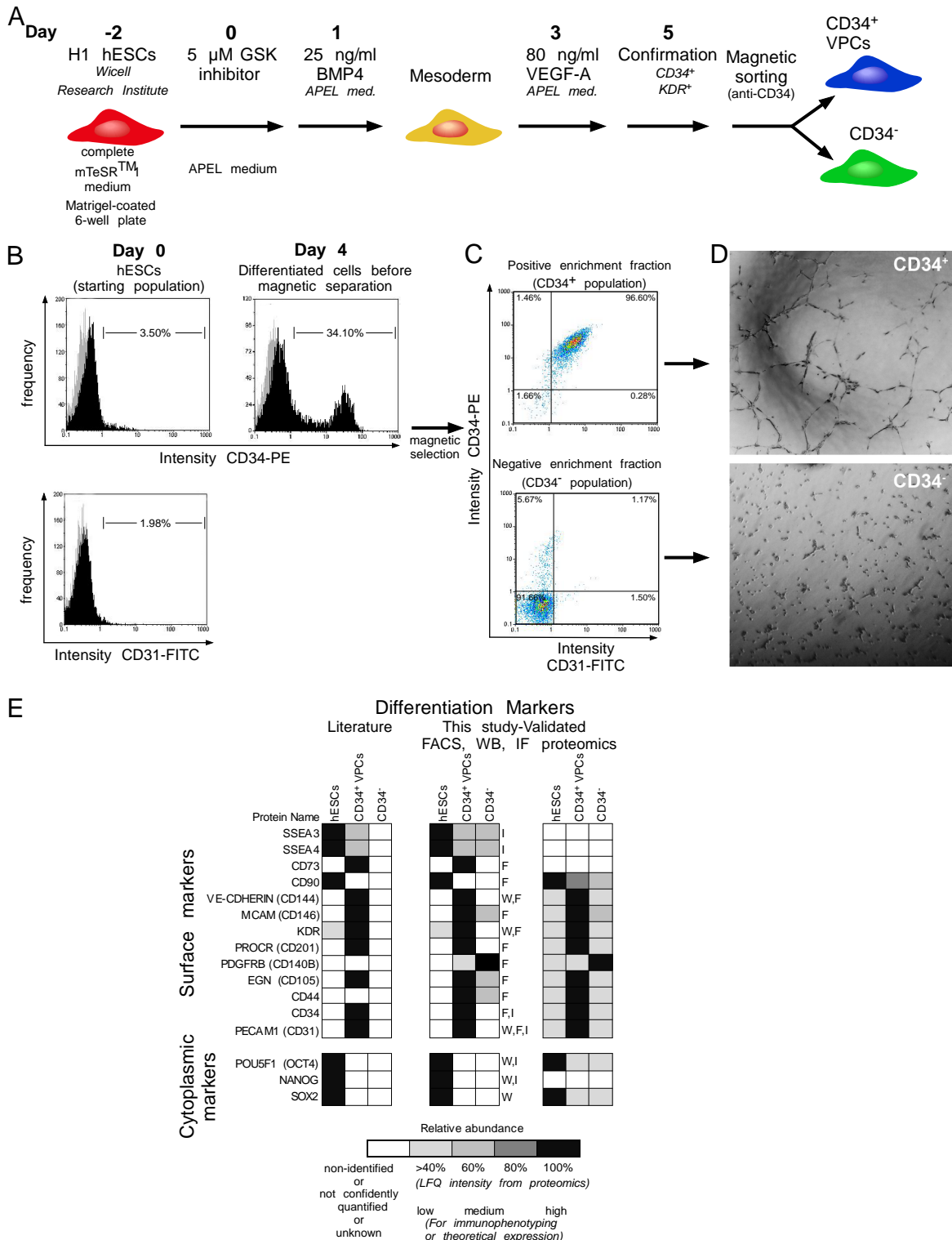Supporting files - statistical analysis scripts.zip: R scripts for the calculation of p-values and adjusted p-values.

for TOC only

Figure 1. Tsolis et al., 2016
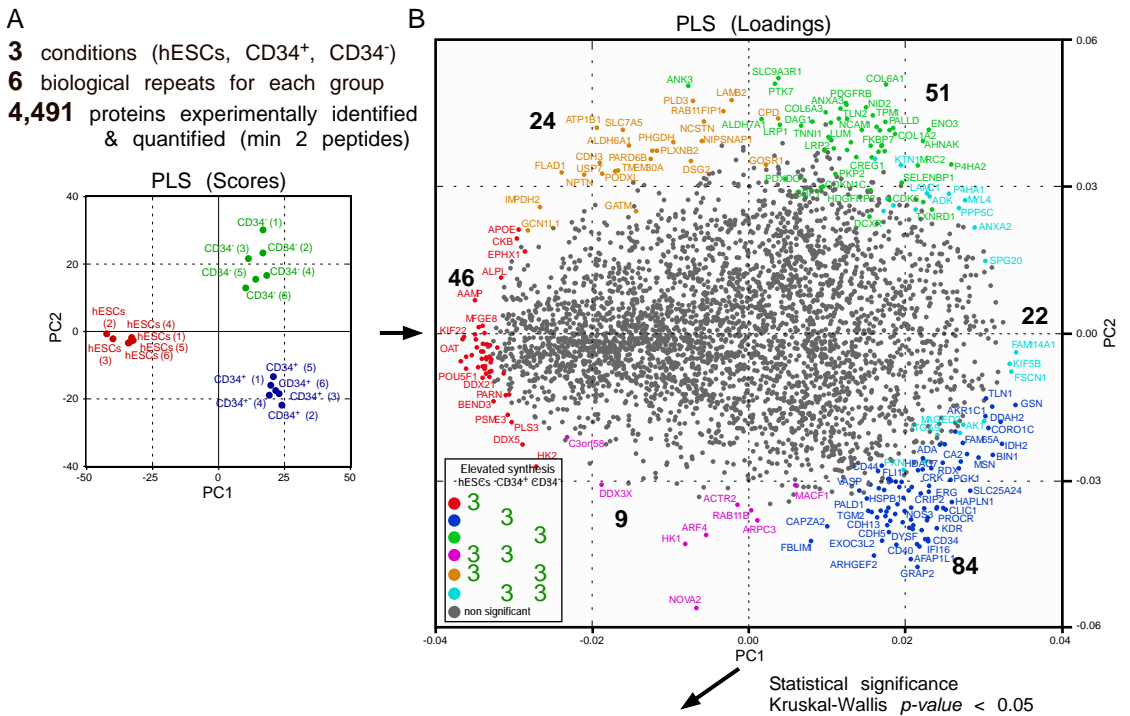
Figure 2. Tsolis et al., 2016

Journal of Proteome Research

Figure 3. Tsolis et al., 2016

Selected proteins of functional interest
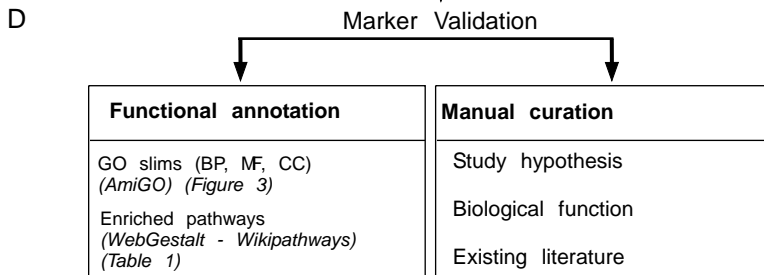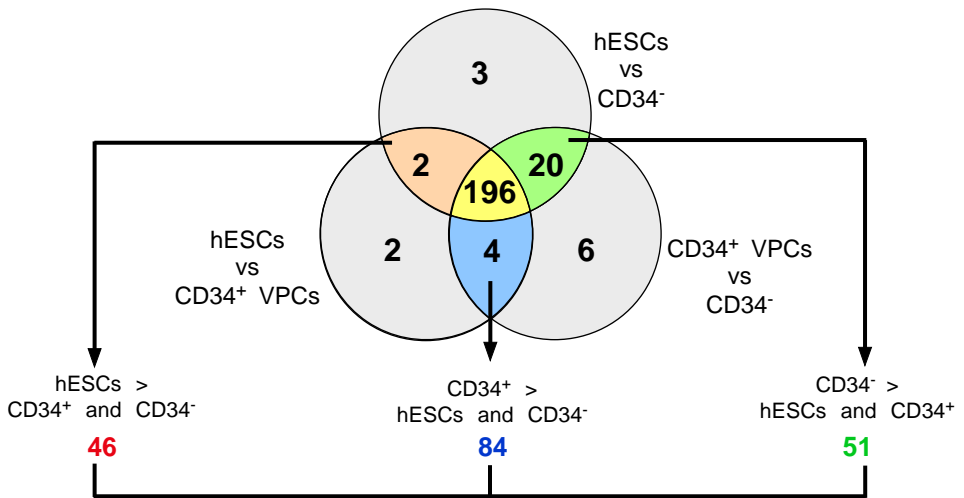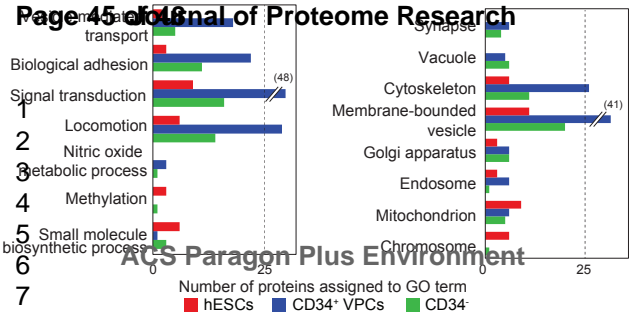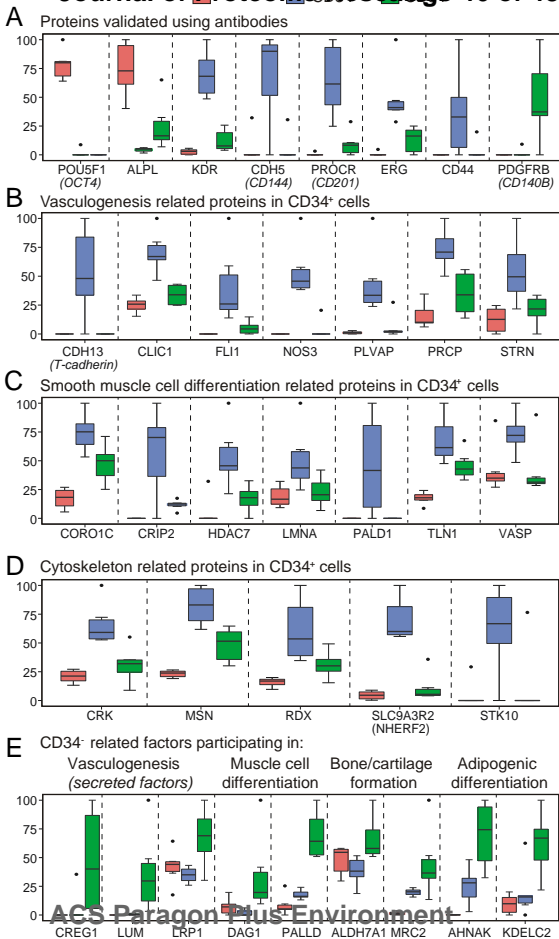
Figure 4. Tsolis et al., 2016

Figure 5. Tsolis et al., 2016

73x40mm (300 x 300 DPI)