Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Cross Study Bulk RNA Sequencing Analysis of Vascular Associated States in Engineered Vascular Systems

Jianhua Zhang

jianhuazt56@outlook.com

Nanchang University https://orcid.org/0009-0009-6422-6993

Research Article

Keywords: vascularized organoid, vascular maturity, vascular smooth muscle cell, mural cell, perfusion, bulk RNA sequencing, cross-study transcriptomic evaluation

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11537048/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Cross Study Bulk RNA Sequencing Analysis of Vascular Associated States in Engineered Vascular Systems

A secondary analysis of publicly available human datasets

Author: Jianhua Zhang (first and corresponding author); ORCID: 0009-0009-6422-6993

Affiliation: School of Pharmacy, Nanchang University, No. 1299 Xuefu Avenue, Honggutan District, Nanchang, Jiangxi 330006, China (Qianhu Campus)

Corresponding author: Jianhua Zhang; Jianhuazt56@outlook.com

Abstract

Background: Vascularization does not necessarily indicate a mature vascular wall. Endothelial cells or vessel-like networks show that vascular-associated structures are present, whereas a stable, functionally mature wall must be assessed using multiple features, including mural-cell coverage, basement membrane, contractile machinery, and functional responses. Because vascularization strategies and outcome measures differ among engineered models, this study evaluates transcriptomic evidence across separate biological dimensions rather than treating vascularization as a single endpoint.

Methods: We performed an exploratory secondary analysis of three publicly available human bulk RNA-sequencing datasets with biological replicates (38 biological libraries in total). The datasets covered vascularized mesenchymal organoids, vascular-network organoids cultured under static or flow conditions, and vascular smooth muscle cells (VSMCs) cultured under two-dimensional or three-dimensional conditions. The core analysis used 11 complementary gene modules to characterize endothelial identity, transcriptional features associated with pericyte-like mural cells, contractile maturity, synthetic remodeling, and flow- and barrier-related programs. Statistical inference used the biological library as the unit, and each dataset retained its original experimental design. CAMERA and expression-matched random-module analyses were applied to the two organoid datasets; Gene Ontology enrichment was limited to three primary comparisons. GSE192986 served as an independent VSMC state reference, with a sensitivity analysis excluding ACTA2.

Results: In the day-14 organoid-versus-spheroid comparison, vascularized mesenchymal organoids showed higher scores in all seven vascular-associated modules (FDR < 0.05 for each); CAMERA supported all seven, although the comparison also differed in scaffold, assembly method, and vascularization conditions. At day 60, several vascular-associated transcriptional signals were lower and hypoxia-related transcription was higher. In a separate vascular-network organoid dataset, high flow reduced hypoxia and inflammatory activation, did not significantly increase endothelial or flow-response module scores, and was associated with a lower contractile-maturity score (FDR = 0.00350). In the independent VSMC dataset, three-dimensional culture increased hypoxia- and inflammation-related expression and reduced proliferation-related expression; contractile maturity did not meet the FDR-adjusted significance threshold.

Conclusions: In these datasets, vascularization-related conditions, culture duration, three-dimensional culture, and flow corresponded to distinct transcriptomic states. Their effects on endothelial, mural, contractile, barrier, and stress-response features should be assessed separately. The findings support a transcriptomic framework for organizing evidence about engineered vascular systems, but they do not by themselves establish vessel-wall structure, perfusion, permeability, or vasomotor function.

Keywords

vascularized organoid; vascular maturity; vascular smooth muscle cell; mural cell; perfusion; bulk RNA sequencing; cross-study transcriptomic evaluation

Introduction

Human organoids reproduce selected features of tissue development, architecture, and cell-cell interactions, but the absence of a functional vascular system can limit their growth and maturation. Inadequate oxygen and nutrient delivery and reduced waste removal may lead to hypoxia or necrosis in organoid cores and leave surrounding parenchymal cells immature. To address this limitation, researchers have introduced vascular components through co-culture with primary endothelial cells, co-differentiation of vascular and organ-specific cells, assembly with vascular organoids, implantation into vascularized hosts, and culture in perfused microfluidic devices [1-6].

Most studies assess vascularization by measuring endothelial markers, counting endothelial cells, or observing vessel-like networks and lumens. These measures mainly indicate the presence of an endothelial compartment or vessel-like structures. Assessment of a mature vascular wall also requires examining whether pericytes, vascular smooth muscle cells, basement membrane, and extracellular matrix form an organized spatial arrangement. Single-cell studies of vascularized organoids have reported differences in vascular-like cell populations, endothelial maturity, and cellular organization among models [7]. In native vessels, these components contribute to vessel stability, barrier properties, vasomotor function, mechanosensing, inflammation, and tissue remodeling [6,8-10]. Endothelial markers indicate that vascular-associated cells or structures are present; a mature vascular unit also requires evidence of organized mural cells and contractile machinery.

This distinction is particularly important in disease modeling. Vascular smooth muscle cells do not remain in a single state. In response to injury or disease, mature contractile cells can shift toward synthetic, proliferative, migratory, inflammatory, or matrix-producing states [10]. This plasticity contributes to atherosclerosis, restenosis, occlusive vascular disease, and pathological vascular remodeling. Vascularized organoids used for mechanistic studies or drug testing therefore need an appropriate baseline vascular-wall state and the ability to transition to states relevant to disease.

Public transcriptomic datasets cover a range of vascularization strategies, organ types, culture durations, and perfusion conditions. Model-specific studies and integrative analyses show that different approaches alter vascular-like cell populations, endothelial maturity, and cellular organization [7-9]. Evaluation should therefore distinguish endothelial vascularization, mural-cell features, contractile vascular-wall maturity, and pathological remodeling. Because studies differ in cell source, organ type, culture duration, and experimental design, merging their gene-expression data would combine these differences in a single result. A more appropriate approach is to retain each study's experimental design and compare conditions within each study at the level of biological samples.

This study evaluated whether a cross-study bulk-transcriptomic framework could separate vascular-associated dimensions in engineered vascular systems. We asked whether vascularized organoids show coordinated transcriptional signals across multiple dimensions, how prolonged culture and flow affect different vascular features, and whether three-dimensional VSMC data provide an independent reference for contractile state. Each dataset was analyzed separately according to its original experimental design. The analysis included three human bulk RNA-sequencing datasets, each with biological replicates. Transcriptomic findings describe molecular features and should be considered alongside spatial, structural, and functional measurements when assessing vascular-wall maturity and function.

Results

Three public datasets were used to assess vascular-associated transcriptomic states

We analyzed the three studies as independent datasets and retained each study's normalized gene-expression data (Figure 1). The verified data comprised 38 biological libraries: GSE297746 included 20 libraries from vascularized mesenchymal organoids collected at different culture

times and day-14 spheroids [5], supporting vascularization and longitudinal comparisons; GSE295400 included 12 libraries from static, low-flow, high-flow, and HUVEC vascular-network conditions [18], supporting assessment of flow; and GSE192986 included six libraries from human VSMCs cultured under two-dimensional or three-dimensional conditions [19], providing a reference for contractile and remodeling states. Each condition had three or four biological replicates.

We used 11 gene modules to describe endothelial identity, mural integration, contractile maturity, synthetic remodeling, flow mechanotransduction, barrier identity, basement membrane, arterial identity, hypoxia response, inflammatory activation, and proliferation (Supplementary Table 2). The flow-mechanotransduction module is referred to below as the flow-response module. The modules overlap biologically and in gene composition; Supplementary Table 9 lists seven module pairs sharing nine genes. We treated the modules as related statistical endpoints rather than independent tests. The mural-integration module captures transcriptomic features associated with pericyte-like mural cells in microvascular walls; it does not directly measure their spatial coverage. Contractile maturity represents the contractile apparatus of differentiated smooth-muscle cells in larger-vessel walls. The two modules address different aspects of the vascular wall.

Day-14 organoids show higher vascular-associated transcriptional signals

We first compared day-14 organoids with day-14 spheroids in GSE297746. Relative to spheroids, organoids had higher scores for endothelial identity, mural integration, and contractile maturity, with module-score differences of 0.948 (95% CI, 0.509 to 1.386; FDR = 0.000528), 0.489 (0.168 to 0.810; FDR = 0.00692), and 0.927 (0.662 to 1.191; FDR = 1.49×10^{-6}), respectively. Scores were also higher for flow response, barrier identity, basement membrane, and arterial identity, with effect estimates of 0.682 (95% CI, 0.324 to 1.040; FDR = 0.00113), 0.761 (0.395 to 1.126; FDR = 0.000539), 0.834 (0.442 to 1.226; FDR = 0.000528), and 0.859 (0.539 to 1.180; FDR = 5.67×10^{-5}), respectively. This comparison also differed in Matrigel scaffold, assembly method, and vascularization conditions, so the observed differences cannot be attributed to vascularization alone. The bulk RNA-seq findings do not establish perfusion or vasomotor function, which require independent structural and functional assessment.

The three analytical approaches gave broadly consistent results (Supplementary Figure S1). CAMERA, which accounts for inter-gene correlation, supported all seven dimensions

(Supplementary Table 4). After multiple-testing correction, the expression-matched random-module analysis supported six dimensions (empirical FDR, 0.00275-0.0128); the mural-integration module had an empirical FDR of 0.113. In the day-14 organoid-versus-spheroid comparison, Gene Ontology enrichment also identified angiogenesis (GO:0001525; NES = 2.49, FDR = 8.55×10^{-15}), blood-vessel development (GO:0001568; NES = 2.06, FDR = 0.00145), blood-vessel morphogenesis (GO:0048514; NES = 2.14, FDR = 0.00124), and extracellular-matrix organization (GO:0030198; NES = 2.39, FDR = 1.26×10^{-8}).

Vascular-associated transcriptional signals decline with prolonged culture

Culture-time comparisons in GSE297746 showed distinct module-specific patterns rather than a uniform trajectory of vascular-associated transcription (Figure 2A). Compared with day 14, day-60 organoids had lower scores in all seven vascular-associated modules. Effect estimates for endothelial identity, mural integration, contractile maturity, flow response, barrier identity, basement membrane, and arterial identity were -1.469, -1.127, -1.084, -0.824, -0.702, -1.677, and -1.651, respectively; the corresponding FDR values were 6.41×10^{-7}, 3.56×10^{-7}, 2.86×10^{-8}, 0.000105, 0.000696, 2.56×10^{-8}, and 1.34×10^{-9}. The 95% confidence intervals are provided in Supplementary Table 3. Hypoxia response increased (0.812, 95% CI, 0.555 to 1.068; FDR = 1.16×10^{-6}), while proliferation-related transcription decreased (-1.290, 95% CI, -1.644 to -0.936; FDR = 2.32×10^{-7}). Synthetic remodeling decreased only in the primary module test (FDR = 0.0056); neither orthogonal test supported this effect. CAMERA supported decreases in endothelial, mural, contractile, flow-response, basement-membrane, and arterial signals, as well as an increase in hypoxia response. The expression-matched random-module analysis supported decreases in endothelial, mural, contractile, basement-membrane, and arterial signals and an increase in hypoxia response; support for flow-response and barrier effects was weaker in this test. Gene-level enrichment also showed increased cellular response to hypoxia (GO:0071456; NES = 1.87, FDR = 0.00293), decreased extracellular-matrix organization (NES = -1.78, FDR = 0.0130), and decreased angiogenesis (NES = -1.48, FDR = 0.0393). In a descriptive analysis, culture day was modeled as a continuous predictor across the 16 organoid libraries. Eight modules showed significant negative trends: endothelial identity, mural integration, flow response, barrier identity, basement membrane, arterial identity, inflammation, and proliferation. Contractile maturity, synthetic remodeling, and hypoxia response did not meet the FDR

threshold (Supplementary Table 7). These module-level trends do not establish a corresponding change in vessel-wall structure or function.

High flow reduces stress-related expression and alters vascular-wall signals

In GSE295400, high flow relative to static culture reduced hypoxia response (-1.163, 95% CI, -1.525 to -0.801; FDR = 0.000166) and inflammatory activation (-1.256, 95% CI, -1.724 to -0.788; FDR = 0.000478). Expression related to contractile maturity, synthetic remodeling, barrier identity, and basement membrane also decreased, with effect estimates of -1.133 (95% CI, -1.716 to -0.549; FDR = 0.00350), -0.479 (-0.731 to -0.226; FDR = 0.00350), -0.580 (-1.018 to -0.141; FDR = 0.0256), and -0.621 (-0.954 to -0.288; FDR = 0.00350), respectively. Proliferation increased (1.124, 95% CI, 0.222 to 2.025; FDR = 0.0297). Neither endothelial identity nor the flow-response module increased significantly. Gene Ontology enrichment also showed reduced response to hypoxia (GO:0001666; NES = -1.97, FDR = 5.11×10^{-5}) and inflammatory response (GO:0006954; NES = -1.75, FDR = 3.64×10^{-4}).

Three-dimensional VSMC culture shows specific cell-state changes

GSE192986 provided an independent reference dataset for human VSMCs (Figure 3). Compared with two-dimensional culture, three-dimensional culture increased hypoxia response (module-score difference, 1.168; FDR = 0.00545) and inflammatory activation (1.030; FDR = 0.0119) and reduced proliferation (-1.440; FDR = 0.000877). Contractile maturity showed a downward trend (-0.694; nominal P = 0.0418; FDR = 0.115). After ACTA2 was excluded, the effect remained negative but was less strongly supported (-0.619; nominal P = 0.0673; FDR = 0.185). CAMERA, expression-matched random-module, and Gene Ontology analyses were not performed for this dataset. Complete 95% confidence intervals are provided in Supplementary Table 3.

Multiple analytical methods support the primary transcriptomic findings

Across methods, day-14 organoids had higher endothelial, contractile, flow-response, barrier, basement-membrane, and arterial signals than spheroids; the mural-integration module showed the weakest support in the matched-null analysis. For day 60 versus day 14, endothelial, mural, contractile, basement-membrane, and arterial signals were lower, while hypoxia-related signals were higher, with broad agreement among the three methods; support for flow-response and barrier effects was weaker. For high flow versus static culture, contractile-, hypoxia-, and

inflammation-related signals decreased and proliferation increased, supported by both orthogonal methods. Results with less consistent support across methods are reported as dimension-specific observations. The main conclusions focus on findings with broadly consistent direction across methods.

Discussion

The main finding is that the tested vascularization-related conditions, culture duration, high flow, and three-dimensional VSMC culture corresponded to distinct transcriptomic states. In GSE297746, day-14 vascularized mesenchymal organoids had higher scores for endothelial, mural, contractile, barrier, basement-membrane, arterial, and flow-response modules than day-14 spheroids. CAMERA supported all seven effects, and the expression-matched random-module analysis supported six; the mural-integration module had weaker support in that null analysis. Because the organoid-versus-spheroid comparison also changed scaffold and assembly conditions, these results should be interpreted as a composite condition contrast rather than an isolated vascularization effect. Lumen patency, continuous mural coverage, vasomotor function, and physiological barrier function require structural and functional assessment.

Culture duration did not correspond to a linear change in vascular maturity. Between days 14 and 60, several vascular-associated transcriptional signals declined in the same model, while hypoxia-related transcription increased. CAMERA, the expression-matched random-module analysis, and gene-level enrichment supported these directions. Tissue composition, central necrosis, dilution of bulk signals, and differences in cell survival may contribute to the changes; the mechanisms require histological and functional testing. Studies at different culture times should report vascular-associated dimensions directly and use them to assess maturity.

High flow reduced stress-related expression and was accompanied by decreases in some vascular-wall-associated signals. Hypoxia- and inflammation-related expression decreased, endothelial and flow-response module scores did not increase significantly, and contractile-maturity scores decreased. Flow rate, waveform, duration, matrix mechanics, vascular geometry, and the developmental state of endothelial and mural cells may all affect the response. Applying flow and forming vessels capable of a mature response to flow are distinct biological steps. A perfusion device can improve oxygen delivery, while endothelial specialization, mural-cell

recruitment, and contractile responses still depend on appropriate developmental and tissue conditions.

The independent VSMC reference dataset further supports this interpretation. Three-dimensional culture increased hypoxia- and inflammation-related expression and reduced proliferation-related expression, while contractile maturity did not increase appreciably. The contractile-module effect remained negative after ACTA2 was excluded, although the evidence was weaker, indicating that the direction was not driven solely by this commonly used smooth-muscle marker. Three-dimensional culture can alter cell phenotype, the extracellular environment, and stress responses. Contractile maturity should therefore be assessed using the coordinated expression of multiple contractile genes and structural or functional evidence. In vascularized organoids, ACTA2-positive cells or signals around vessel-like structures should also be considered alongside MYH11, CNN1, TAGLN, MYOCD, and SMTN expression and direct structural or functional measurements.

These findings support a staged evaluation framework. First, establish whether endothelial cells or vessel-like structures are present. Next, assess mural-cell coverage, contractile vascular-wall features, basement membrane, and barrier features. Then test responses to flow or disease-related perturbations. The relevant measures depend on the intended use: an angiogenesis model may prioritize endothelial sprouting, whereas a model for vasomotor pharmacology requires mature mural coverage and stimulus-responsive contraction. Reporting each dimension separately clarifies the basis for model selection and prevents a strong endothelial signal from obscuring deficits in other dimensions.

This study has several limitations. Bulk RNA sequencing captures tissue-level transcriptional signals; spatial proximity between endothelial and mural cells, continuity of mural coverage, and the relative contributions of cell abundance and cell state require single-cell RNA sequencing, spatial transcriptomics, and histology. The day-14 organoid-versus-spheroid comparison also differs in Matrigel, assembly method, and vascularization conditions and therefore does not isolate one causal factor. The day-60 versus day-14 comparison may be affected by tissue composition, central necrosis, dilution of bulk signals, or differences in cell survival. The high-flow comparison may also reflect mural-cell loss or changes in cell composition under shear. The three datasets represent different biological systems, so their within-study effects are reported separately and interpreted as complementary evidence rather than pooled estimates. Public

metadata do not fully establish donor, clone, batch, or pairing structures; parallel labels such as 2D_1 and 3D_1 were treated as independent libraries. A group-based or paired sensitivity analysis could be conducted if complete mappings become available. The GSE295400 organoid conditions come from a single WTC cell line. GSE192986 is a short-term VSMC state-reference experiment that complements the contractile and remodeling findings; an independent dataset or experiment is needed to represent native arterial states. The analysis is intended to generate and organize transcriptomic evidence, not to establish functional vascular maturity without independent validation.

Future validation can proceed across structural, functional, and disease-model dimensions. First, sample-traceable transcriptomic data can be combined with quantitative mural coverage, basement-membrane continuity, and lumen patency. Next, studies can measure tracer permeability and flow and test contraction or dilation after vasoactive stimulation. Finally, they can determine whether a model can first establish a mature vascular unit and then transition reproducibly to a state relevant to disease. Module scores provide molecular evidence of vascular-associated biology; a fully functional vascular circulation also requires permeability, perfusion, and vasomotor testing.

In summary, vascularization and vascular-wall maturity are related but distinct dimensions. Prolonged culture, high flow, three-dimensional organization, vessel-like morphology, and a single smooth-muscle marker each reflect only some aspects of the vascular system. Maturity should be evaluated using endothelial, mural, contractile, barrier, and functional-response measures together. Within the engineered vascular systems covered here, we retained each study's experimental design and used biological libraries as the units of inference to organize a tiered evaluation from endothelial features and mural-associated transcriptional signals to contractile and functional responses. This framework is a hypothesis-generating aid for model selection and engineering, not a substitute for structural or functional validation.

Methods

Study design and dataset eligibility

This exploratory secondary analysis used three publicly available human bulk RNA-sequencing datasets comprising 38 biological libraries. It was not prospectively registered; the gene modules and primary contrasts were defined for this analysis and are not prospective confirmatory

endpoints. Datasets were eligible if they profiled a human vascularized organoid, engineered vascular network, or human VSMC model; included an interpretable comparator; provided gene-level counts or an author-supplied matrix containing counts; and included at least three biological libraries per tested condition. Datasets without biological replication were not used for inferential statistical analysis.

GSE297746 comprised 20 libraries (four replicates per condition) from free-floating vascularized mesenchymal organoids cultured for 3, 14, 28, or 60 days and day-14 HUVEC-mesenchymal stromal cell spheroids without a Matrigel scaffold [5]. GSE295400 comprised 12 libraries (three replicates per condition) from pluripotent-stem-cell-derived vascular-network organoids cultured under static, low-flow, or high-flow conditions and an HUVEC-network reference [18]. The organoid conditions in this study came from a single WTC pluripotent stem-cell line, and the findings primarily apply to this line. GSE192986 comprised six libraries (three replicates per condition) from primary human VSMCs cultured for four days under two-dimensional conditions or in three-dimensional polyglycolic-acid scaffolds [19]. The datasets were used, respectively, for vascularization and longitudinal comparisons, flow-response assessment, and VSMC state reference. Low-flow conditions in GSE295400 were treated as an additional flow condition and are reported in the complete result tables; the primary results focus on high flow versus static culture.

Data acquisition and provenance

Series metadata and processed gene-level files were obtained from the Gene Expression Omnibus. File sources, integrity checks, and SHA-256 checksums were recorded to make the secondary-analysis inputs traceable and support reproducibility. The GSE297746 and GSE295400 supplemental archives passed tar integrity checks, and the GSE192986 expression matrix passed gzip integrity testing. The analysis used per-sample count files for GSE297746, the NumReads columns for GSE295400, and six count.* columns for GSE192986. Although the GSE192986 file name contains FPKM, the analysis used its integer-valued count.* columns as edgeR input.

Gene annotation, normalization, and module scores

Ensembl version suffixes were removed where present. Ensembl identifiers were mapped to HGNC gene symbols with org.Hs.eg.db; counts for identifiers mapping to the same symbol were

summed. GSE192986 was collapsed by the deposited gene_name field. Within each dataset, trimmed mean of M-values normalization factors were estimated with edgeR, and log2 counts per million were calculated with a prior count of 2 [11,12]. These steps harmonized gene identifiers and sequencing depth so that module scores could be compared within each study.

The 11 modules used in this analysis represented endothelial identity, mural integration, contractile maturity, synthetic remodeling, flow mechanotransduction, barrier identity, basement membrane, arterial identity, hypoxia response, inflammatory activation, and proliferation. Complete module definitions, applicable analysis levels, audit fields, and gene coverage by dataset are reported in Supplementary Table 2; shared genes are listed in Supplementary Table 9. The Primary set column indicates whether a gene belongs to the primary module set; all 114 genes belonged to this set, which did not confer statistical weight. The mural-integration module captures transcriptomic features associated with pericyte-like mural cells in microvascular walls; it does not directly measure their spatial coverage. Contractile maturity represents the contractile apparatus of differentiated smooth-muscle cells in larger-vessel walls. The analysis records do not establish whether the modules were fixed before data inspection; missing source citations and definition dates are identified as unrecorded in Supplementary Table 2. These modules and contrasts were selected for an exploratory secondary analysis and should not be interpreted as prospectively registered endpoints. Within each dataset, logCPM values for each detected gene were standardized across all biological libraries. Each sample's module score was the mean of the gene-wise z scores for its detected module genes and was calculated only when at least three module genes were detected. Because standardization was performed within each study, scores were compared only across conditions in the same dataset; absolute scores were not pooled across datasets.

Sample-level module inference

Biological libraries were the units of statistical inference. This avoids treating genes or inferred cell counts as replicates. For each dataset, group means were estimated with a no-intercept design matrix in limma [13]. Contrasts were fitted with contrasts.fit, followed by robust empirical Bayes moderation with eBayes(robust = TRUE). Effect estimates are differences in mean module score and are reported with two-sided 95% confidence intervals that were not adjusted for multiple comparisons. The primary comparisons were day-14 organoid versus day-14 spheroid and day-60 versus day-14 organoid in GSE297746; high flow versus static

culture in GSE295400; and three-dimensional versus two-dimensional VSMC culture in GSE192986. The first three were the primary organoid comparisons; the VSMC comparison provided an independent reference. These labels describe the role of the comparisons in this analysis and do not indicate prospective registration. Other time and flow comparisons are reported in the complete result tables. Benjamini-Hochberg correction was applied across the 11 modules within each comparison, with no joint FDR correction across comparisons. Complete results are provided in Supplementary Table 3.

Gene-level differential expression and pathway enrichment

Gene-level count analyses were performed separately for each dataset with edgeR [11,14]. This model was used to test gene-level changes within each study and provide ranked statistics for Gene Ontology enrichment. Lowly expressed genes were removed with filterByExpr, followed by trimmed mean of M-values normalization and robust dispersion estimation. Quasi-likelihood models were fitted with glmQLFit(robust = TRUE). Each contrast included in the analysis was tested with glmQLFTest, and gene-level P values were adjusted using the Benjamini-Hochberg method.

Gene Ontology Biological Process sets were assembled from org.Hs.eg.db and GO.db [17] and restricted to sets containing 15-500 detected genes in the corresponding dataset. Genes were ranked by the square root of the edgeR quasi-likelihood F statistic, with direction assigned according to the sign of the contrast coefficient (log-fold change). Enrichment was tested with fgseaMultilevel using eps = 0 [16] to complement the module results and identify related biological processes. The analysis covered only the three primary comparisons: day-14 organoid versus spheroid and day-60 versus day-14 organoid in GSE297746, and high flow versus static culture in GSE295400. Multiple-testing correction was applied across all tested Gene Ontology sets within each dataset and comparison. No GO enrichment analysis was performed for GSE192986. Complete test results and leading-edge genes are reported in Supplementary Table 5.

Competitive and expression-matched robustness tests

CAMERA was used to re-test the modules included in this analysis and assess the effect of inter-gene correlation on the main findings [15]. CAMERA and expression-matched random-module analyses were applied only to the GSE297746 and GSE295400 organoid datasets, not to the

GSE192986 VSMC reference. For applicable datasets, genes passing filterByExpr were TMM-normalized and transformed with voom. CAMERA used the original no-intercept group design and the same contrast vectors as the primary analysis, with inter.gene.cor = 0.01. Modules with fewer than three detected genes were excluded. CAMERA false-discovery rates were calculated across modules within each comparison.

To assess whether arbitrary genes with similar expression abundance could reproduce an observed module effect, genes with nonzero cross-sample variance were ranked by mean normalized expression and divided into 20 mean-expression bins. For each module and comparison, 2,000 random modules were sampled without replacement while preserving the number of target-module genes in each bin. Genes belonging to any of the 11 analysis modules were excluded from the sampling pool. The empirical two-sided P value was the proportion of random effects whose absolute deviation from the null-distribution mean exceeded the corresponding absolute deviation of the observed effect; a plus-one correction was applied. Empirical P values were adjusted across modules within each dataset and comparison using the Benjamini-Hochberg method. The random seed was 287254.

As a descriptive supplement, culture day was modeled as a continuous predictor in an unpaired linear model across the 16 GSE297746 organoid libraries. The estimated daily trends for each module are reported in Supplementary Table 7. Public metadata did not provide a verifiable longitudinal pairing map, so this model provides a descriptive analysis; paired or blocked analysis would be a more rigorous approach to estimating time effects.

Sensitivity analysis and cross-study interpretation

All comparisons involving contractile maturity were repeated after excluding ACTA2 to assess whether the findings depended on this commonly used smooth-muscle marker (Supplementary Table 6). Normalization and statistical analyses were performed within each study. Because the three systems, interventions, and biological comparisons had different experimental contexts, cross-study synthesis was qualitative and based on effect directions and statistical evidence within each study. Module scores were interpreted as molecular evidence for the corresponding biological process and considered alongside structural and functional measurements of endothelial-mural proximity, vessel-wall coverage, lumen patency, permeability, perfusion, and vasomotor function.

Ethics statement

This study used only publicly available, de-identified data from the Gene Expression Omnibus (GEO). It involved no new participant recruitment, intervention, specimen collection, or access to identifiable information. No new participant-related procedures were conducted for this secondary analysis. Ethics approval and informed-consent procedures for the original studies are cited from the corresponding publications or GEO records where reported; no inference is made about information that was not reported.

Data and code availability

All primary expression data analyzed in this study are publicly available from the GEO accessions listed in Supplementary Table 1. Supplementary_Tables.xlsx provides Supplementary Tables 1-9, covering datasets, module definitions, module effects, robustness tests, Gene Ontology enrichment, ACTA2 sensitivity, linear time trends, the supplementary-table index, and module overlap. The original GEO expression files are not redistributed.

The analysis code is not included with this preprint and is not deposited in a public repository. No separate public repository DOI is available at this submission stage. The code is not distributed under an open-source license.

The supplementary documentation and derived results are provided under CC BY 4.0.

Declarations

Funding

No specific funding was received for this study.

Competing interests

The author declares no competing interests.

Author contributions

Jianhua Zhang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, and Writing - review and editing.

Use of generative AI

Generative AI tools assisted with coding.

References

1. Homan KA, Gupta N, Kroll KT, et al. Flow-enhanced vascularization and maturation of kidney organoids in vitro. Nat Methods. 2019;16(3):255-262. doi:10.1038/s41592-019-0325-y. PMID:30742039.
2. Cakir B, Xiang Y, Tanaka Y, et al. Engineering of human brain organoids with a functional vascular-like system. Nat Methods. 2019;16(11):1169-1175. doi:10.1038/s41592-019-0586-5. PMID:31591580.
3. Sun XY, Ju XC, Li Y, et al. Generation of vascularized brain organoids to study neurovascular interactions. eLife. 2022;11:e76707. doi:10.7554/eLife.76707. PMID:35506651.
4. Quintard C, Tubbs E, Jonsson G, et al. A microfluidic platform integrating functional vascularized organoids-on-chip. Nat Commun. 2024;15(1):1452. doi:10.1038/s41467-024-45710-4. PMID:38365780.
5. de Janon A, Kaza P, Yoshida K, et al. Free-floating long-term vascularized mesenchymal organoids. iScience. 2026;29(1):114548. doi:10.1016/j.isci.2025.114548. PMID:41568151.
6. Naderi-Meshkin H, Cornelius VA, Eleftheriadou M, Potel KN, Setyaningsih WAW, Margariti A. Vascular organoids: unveiling advantages, applications, challenges, and disease modelling strategies. Stem Cell Res Ther. 2023;14(1):292. doi:10.1186/s13287-023-03521-2. PMID:37817281.
7. Sato Y, Asahi T, Kataoka K. Integrative single-cell RNA-seq analysis of vascularized cerebral organoids. BMC Biol. 2023;21(1):245. doi:10.1186/s12915-023-01711-1. PMID:37940920.
8. Tan SY, Jing Q, Leung Z, Xu Y, Cheng LKW, Tam SST, Wu AR. Transcriptomic analysis of 3D vasculature-on-a-chip reveals paracrine factors affecting vasculature growth and maturation. Lab Chip. 2022;22(20):3885-3897. doi:10.1039/d2lc00570k. PMID:36093896.
9. Koning M, Dumas SJ, Meta E, et al. Single cell transcriptomics of human kidney organoid endothelium reveals vessel growth processes and arterial maturation upon transplantation. NPJ Regen Med. 2025;10(1):32. doi:10.1038/s41536-025-00418-x. PMID:40595660.
10. Basatemur GL, Jorgensen HF, Clarke MCH, Bennett MR, Mallat Z. Vascular smooth muscle cells in atherosclerosis. Nat Rev Cardiol. 2019;16(12):727-744. doi:10.1038/s41569-019-0227-9. PMID:31243391.
11. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-140. doi:10.1093/bioinformatics/btp616. PMID:19910308.
12. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11(3):R25. doi:10.1186/gb-2010-11-3-r25. PMID:20196867.
13. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47. doi:10.1093/nar/gkv007. PMID:25605792.
14. Chen Y, Lun ATL, Smyth GK. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000Res. 2016;5:1438. doi:10.12688/f1000research.8987.2. PMID:27508061.
15. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res. 2012;40(17):e133. doi:10.1093/nar/gks461. PMID:22638577.

16. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. bioRxiv. Posted February 1, 2021. doi:10.1101/060012.

17. Gene Ontology Consortium. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. 2021;49(D1):D325-D334. doi:10.1093/nar/gkaa1113. PMID:33290552.

18. National Center for Biotechnology Information. GSE295400: Engineering perfusable vascular networks in organoids-on-chip [bulk RNA-seq]. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE295400. Accessed October 11, 2026.

19. National Center for Biotechnology Information. GSE192986: Transcriptome analysis with human vascular smooth muscle cells under 2D and 3D cell culture conditions. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE192986. Accessed October 11, 2026.

478 Figure 1



479

480 **Figure 1** Study design and analysis workflow. (A) Evaluation dimensions include endothelial presence, mural
481 integration, vascular-wall maturity, and functional responses. Each dimension is assessed separately, and no
482 single marker is interpreted in isolation when evaluating vascular-wall maturity. (B) The three transcriptomic
483 datasets and their roles in the analysis; sample counts refer to the number of biological libraries included in
484 each public study. (C) Analysis workflow. Each dataset was processed and tested within study, with biological
485 libraries as the units of inference. CAMERA and expression-matched random-module analyses were
486 performed for the two organoid datasets; Gene Ontology enrichment was performed for the three primary
487 comparisons. ACTA2-excluded sensitivity and linear time-trend analyses are reported in Supplementary
488 Tables 6 and 7.

Figure 2

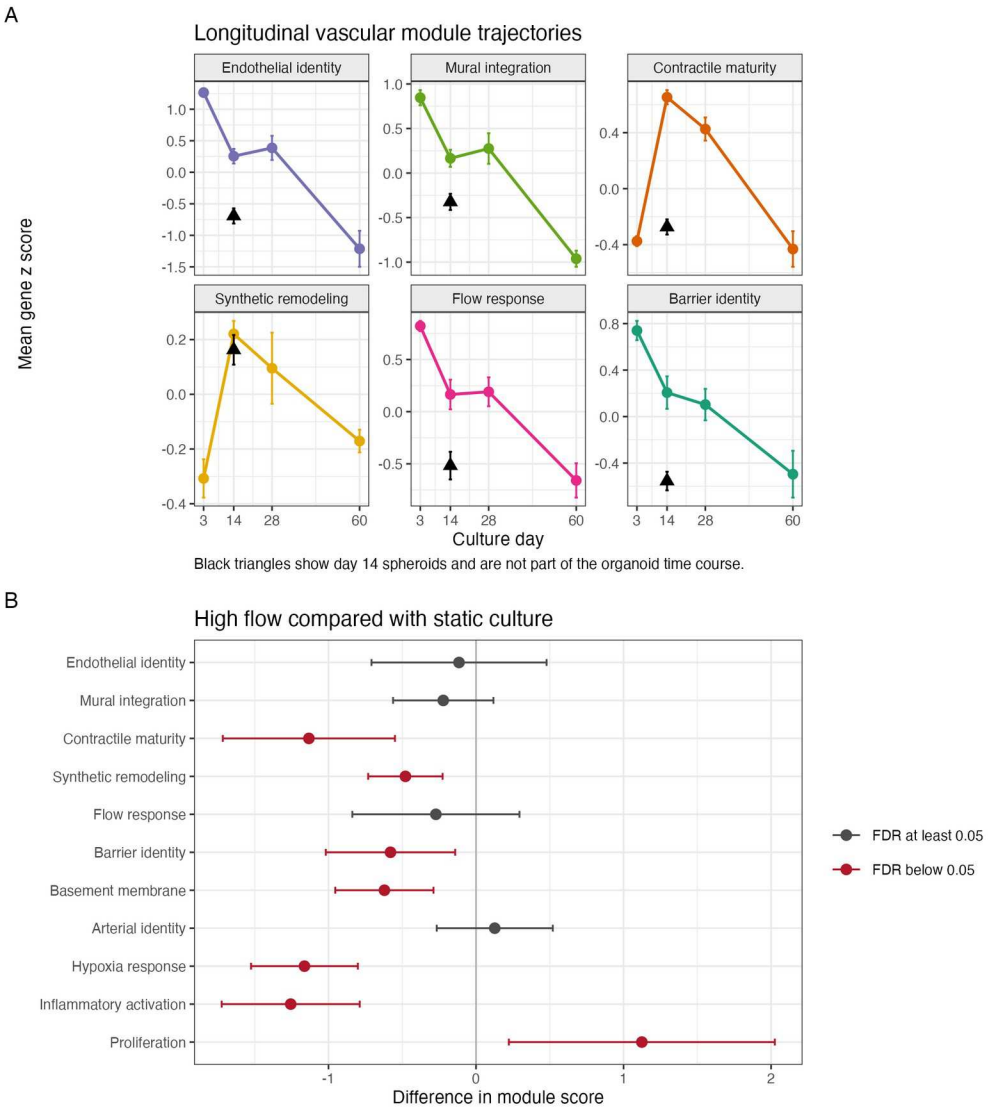


Figure 2 Effects of culture duration and high flow on vascular-associated modules. (A) Trajectories of mean gene-wise z scores for six modules (endothelial identity, mural integration, contractile maturity, synthetic remodeling, flow response, and barrier identity) across culture days in vascularized mesenchymal organoids from GSE297746. Points show group means and error bars show standard errors. The black triangle marks the day-14 spheroid group, which is not part of the organoid time course. (B) Differences in module scores between high-flow and static vascular-network organoids in GSE295400. Points show estimated mean differences and horizontal bars show 95% confidence intervals. Red indicates a Benjamini-Hochberg FDR below 0.05 in the sample-level module analysis.

Figure 3

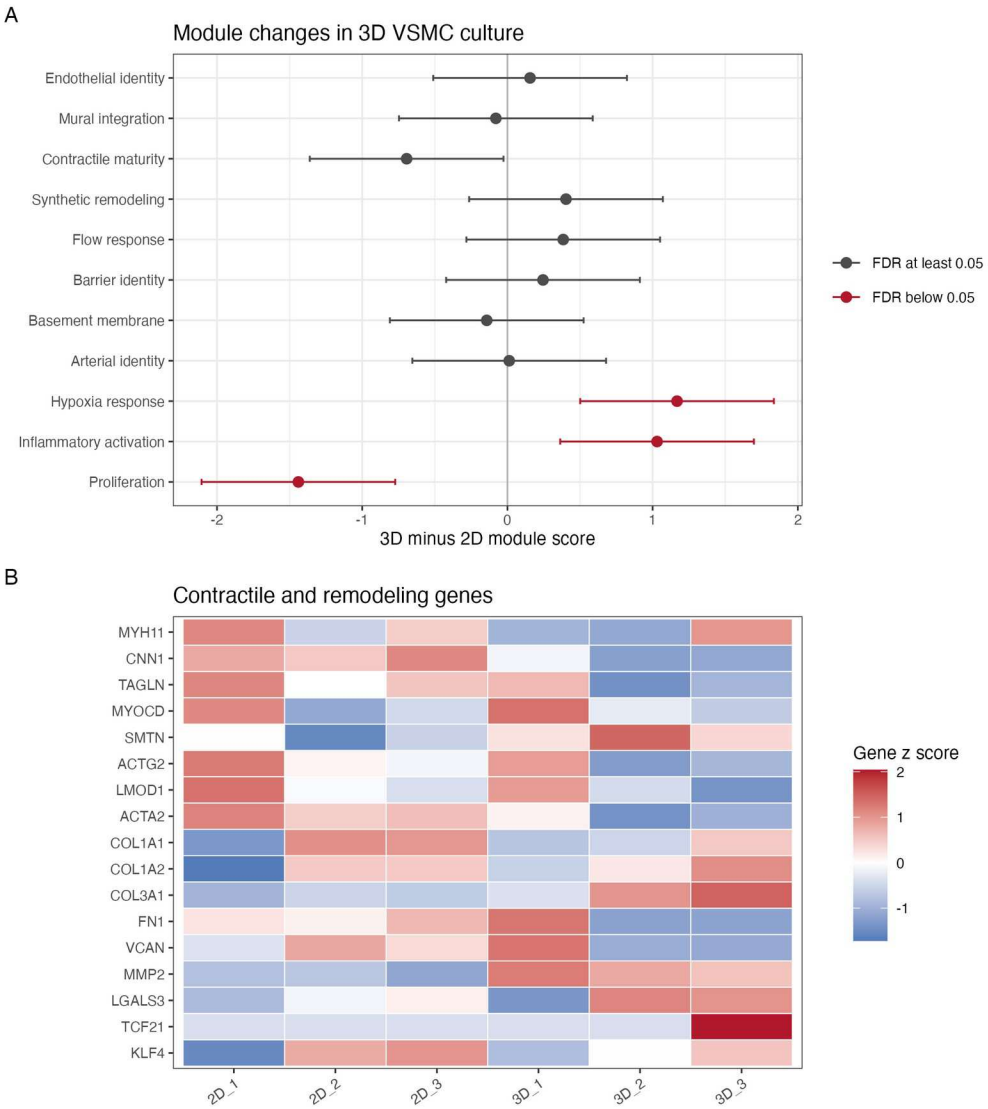
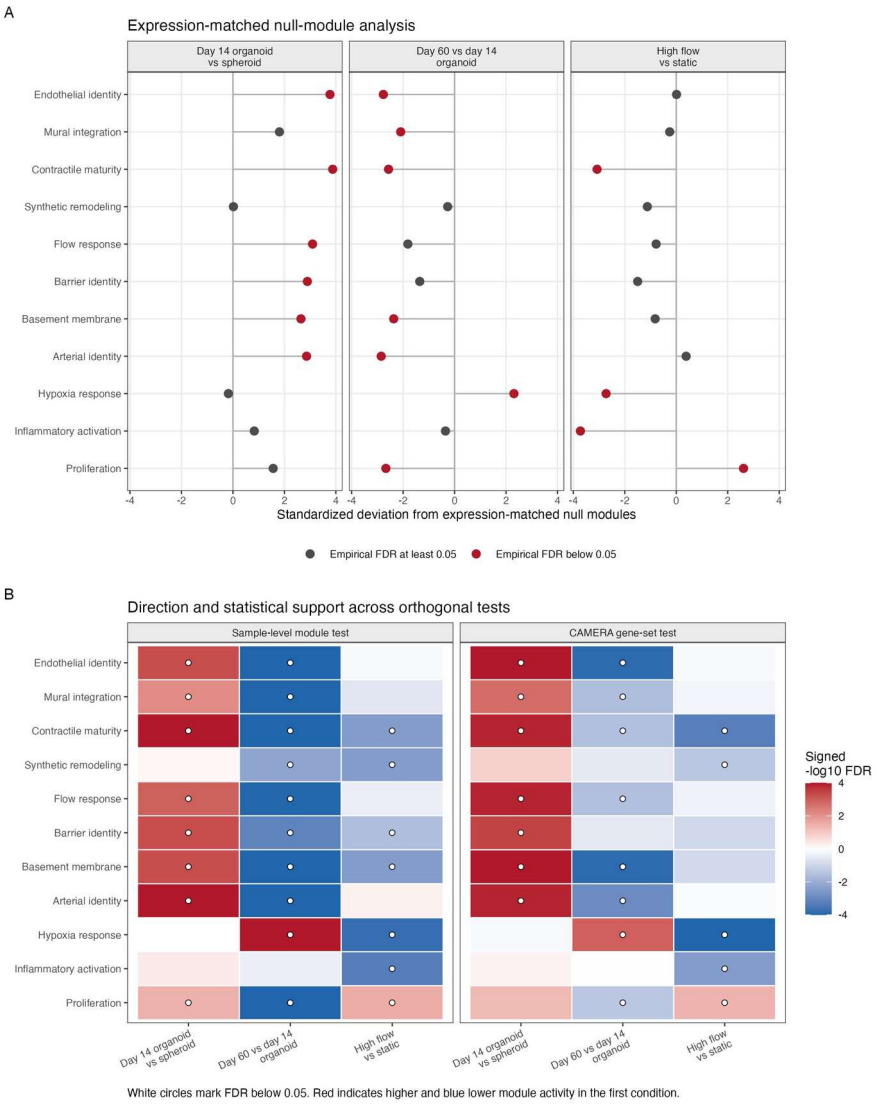


Figure 3 Transcriptomic comparison of VSMCs cultured under 2D and 3D conditions. (A) Differences in module scores between three-dimensional and two-dimensional human VSMC cultures in GSE192986. Points show estimated mean differences and horizontal bars show 95% confidence intervals. Red indicates a Benjamini-Hochberg FDR below 0.05. (B) Gene-wise z scores for selected contractile and remodeling genes across three biological replicates per condition. This experiment describes VSMC phenotypic state and provides an independent reference for the organoid findings; contractile-related transcription under three-dimensional culture should be interpreted alongside other molecular, structural, and functional evidence.

Supplementary Figure S1



Supplementary Figure S1. Robustness of bulk RNA-sequencing module analyses. (A) For each of the three primary comparisons, the standardized deviation of each observed module effect from the null distribution generated by 2,000 expression-matched random modules is shown. The null modules matched each target module in size and mean-expression-bin composition. Red points indicate empirical FDR < 0.05. (B) Effect direction and statistical support from sample-level module tests and CAMERA competitive gene-set tests. Colors indicate signed -log10 FDR (absolute values capped at 4); red indicates a higher module score in the first condition and blue a lower score. White circles indicate FDR < 0.05.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarytables.xlsx
- SupplementaryFigureS1.pdf