

## P4HA2: A link between tumor-intrinsic hypoxia, partial EMT and collective migration

Vaishali Aggarwal<sup>a</sup>, Sarthak Sahoo<sup>b</sup>, Vera S. Donnenberg<sup>a,c,d,e</sup>, Priyanka Chakraborty<sup>b</sup>, Mohit Kumar Jolly<sup>b</sup>, Shilpa Sant<sup>a,c,d,e,f,\*</sup>

<sup>a</sup> Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh, PA, USA

<sup>b</sup> Centre for BioSystems Science and Engineering, Indian Institute of Science, Bangalore, India

<sup>c</sup> Department of Cardiothoracic Surgery, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA

<sup>d</sup> McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, PA, USA

<sup>e</sup> UPMC-Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA, USA

<sup>f</sup> Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA, USA

### ARTICLE INFO

#### Keywords:

P4HA2  
Tumor-intrinsic hypoxia  
Partial EMT  
Collective migration  
3D microtumor models  
Tumor microenvironment

### ABSTRACT

Epithelial-to-mesenchymal transition (EMT), a well-established phenomenon studied across pan-cancer types, has long been known to be a major player in driving tumor invasion and metastasis. Recent studies have highlighted the importance of partial EMT phenotypes in metastasis. Initially thought as a transitional state between epithelial and mesenchymal phenotypic states, partial EMT state is now widely recognized as a key driver of intra-tumoral heterogeneity and phenotypic plasticity, further accelerating tumor metastasis and therapeutic resistance. However, how tumor microenvironment regulates partial EMT phenotypes remains unclear. We have developed unique size-controlled three-dimensional microtumor models that recapitulate tumor-intrinsic hypoxia and the emergence of collectively migrating cells. In this study, we further interrogate these microtumor models to understand how tumor-intrinsic hypoxia regulates partial EMT and collective migration in hypoxic large microtumors fabricated from T47D breast cancer cells. We compared global gene expression profiles of hypoxic, migratory microtumors to that of non-hypoxic, non-migratory microtumors at early and late time-points. Using our microtumor models, we identified unique gene signatures for tumor-intrinsic hypoxia (early *versus* late), partial EMT and migration (pre-migratory *versus* migratory phenotype). Through differential gene expression analysis between the microtumor models with an overlap of hypoxia, partial EMT and migration signatures, we identified prolyl 4-hydroxylase subunit 2 (P4HA2), a hypoxia responsive gene, as a central regulator common to hypoxia, partial EMT and collective migration. Further, the inhibition of P4HA2 significantly blocked collective migration in hypoxic microtumors. Thus, using the integrated computational-experimental analysis, we identify the key role of P4HA2 in tumor-intrinsic hypoxia-driven partial EMT and collective migration.

### 1. Introduction

Epithelial-to-mesenchymal transition (EMT) during which epithelial

cells (E) lose their cell-cell adhesion properties and transition to a mesenchymal state (M), has long been known to drive tumor invasion and metastasis [1–3]. Historically, EMT was thought to have terminal E

**Abbreviations:** Epithelial, (E); Mesenchymal, (M); Prolyl 4-hydroxylase subunit 2, (P4HA2); Epithelial-to-mesenchymal transition, (EMT); E-cadherin, (E-CAD); Vimentin, (VIM); Neural Cadherin, (NCAD); Tumor microenvironment, (TME); Hypoxia inducible factor 1 alpha, (HIF1a); Cytokeratin, (CTK); Gene Expression Omnibus, (GEO); Gene Set Enrichment Analysis, (GSEA); Molecular Signatures Database, (MsigDB); Single-sample GSEA, (ssGSEA); The Cancer Genome Atlas, (TCGA); 1,4-dihydrophenanthrolin-4-one-3-carboxylic acid, (1,4-DPCA); Prolyl 4-hydroxylase subunit 1, (P4HA1); NUMB endocytic adaptor protein, (NUMB); Extracellular matrix protein 1, (ECM1); Grainyhead like transcription factor 2, (GRHL2); Thrombospondin 1, (THBS1); Protein-protein interaction, (PPI); Cadherin 1, (CDH1); Phenotypic stability factor, (PSF); Ductal carcinoma in situ, (DCIS).

\* Corresponding author. University of Pittsburgh School of Pharmacy Department of Pharmaceutical Sciences Department of Bioengineering UPMC-Hillman Cancer Center McGowan Institute for Regenerative Medicine, 7408 Salk Hall, 3501 Terrace Street, Pittsburgh, PA, 15261, USA.

E-mail address: [shs149@pitt.edu](mailto:shs149@pitt.edu) (S. Sant).

<https://doi.org/10.1016/j.adcanc.2022.100057>

Received 28 February 2022; Received in revised form 24 May 2022; Accepted 25 May 2022

Available online 31 July 2022

2667-3940/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

and M states and earlier studies focused on underlying mechanisms driving EMT [4,5]. Emerging evidence suggests the existence of partial EMT phenotype, a transitional state between the E and M states, wherein the cells retain the epithelial phenotype with gain of mesenchymal phenotype [1,6]. Currently, partial EMT phenotype is characterized by co-expression of epithelial (ECAD, ZO-1) and mesenchymal markers (VIM, SNAI1,  $\beta$ -catenin, ZEB1, FN1, MMP1, ZEB2, NCAD) [7]. The partial EMT phenotypes contribute to tumor heterogeneity, cell cycle regulation, collective migration, metastasis and therapeutic resistance [7]. It is suggested that the E, M and partial EMT subsets of malignant cells are present concurrently during tumor progression driving the intra-tumoral heterogeneity [8,9]. However, the molecular mechanisms that drive partial EMT states still remain understudied.

It is widely accepted that the tumor microenvironment (TME) plays a critical role in tumor progression [10]. TME is shaped by tumor-intrinsic factors (e.g., hypoxia, acidic microenvironment) and tumor-extrinsic factors (extracellular matrix and stromal related factors) [10]. TME consists of heterogeneous cellular composition including malignant cells, endothelial cells, stromal cells, fibroblasts, immune cells and extracellular matrix [10,11]. Of these, tumor-intrinsic hypoxia has been suggested to drive and stabilize partial EMT [8,10,12]. In oesophageal squamous cell carcinoma, malignant cells overexpressed hypoxia inducible factor 1 alpha (HIF1a) and exhibited partial EMT phenotype characterized by co-expression of E-cadherin (E-CAD) and vimentin (VIM) [13]. Acidic microenvironment also contributed to partial EMT in MCF-7 cells, which co-expressed epithelial (ZO-1 and E-CAD) and mesenchymal (VIM, SNAI1 and  $\beta$ -catenin) markers [14]. In addition to tumor-intrinsic factors, tumor-extrinsic factors such as extracellular matrix (fibrinogen or collagen matrix) [15,16], cancer-associated fibroblasts and stromal cells [17,18] and tumor associated macrophages [19,20] have all shown to drive and sustain partial EMT phenotypes. These studies support the collective role of TME in inducing partial EMT tumor states.

To investigate mechanisms by which tumor-intrinsic hypoxia drives partial EMT phenotype, it is imperative to have experimental model systems, which can successfully recapitulate inherent tumor-intrinsic hypoxic microenvironment. Majority of published studies utilize exposure of 2D cell monolayers or three-dimensional (3D) spheroids to hypoxic chambers, where all cells in 2D monolayers or even outer cells in 3D spheroids are exposed to constant hypoxia, in contrast to the dynamic, spatial oxygen gradients that are observed *in vivo* [10]. Our lab has engineered a unique 3D microtumour models, which recapitulate tumor-intrinsic hypoxic microenvironment, solely driven by the size of microtumors without any genetic manipulation or any artificial culture conditions [12,21–23]. A major advantage of microtumor models is the ease and reproducibility of generation of both non-hypoxic, non-migratory microtumors (150  $\mu$ m, control) as well as hypoxic, migratory microtumors (600  $\mu$ m) *in vitro*. We have shown that the migratory phenotype is easily induced in large hypoxic microtumors made of parent non-migratory T47D cells without any external stimulus [21–23]. We have also observed the presence of heterogenous tumor states *i.e.*, E, M and E/M (partial EMT) [23]. Specifically, we have shown that the large hypoxic microtumors gain the expression of mesenchymal marker (VIM) between day 3–6 while maintaining epithelial marker (E-CAD) expression, supporting the acquisition of partial EMT phenotype [23]. Subsequent studies highlighted the co-localization of E-CAD<sup>+</sup> VIM<sup>+</sup> or CD90<sup>+</sup>CD44<sup>+</sup>CTK<sup>+</sup> (partial EMT) tumor cells at the leading edge of the migratory front, further solidifying the role of partial EMT in driving collective migration [23,24].

The current study delineates how tumor-intrinsic hypoxia drives partial EMT and collective migration. To investigate the potential mechanisms underlying tumor-intrinsic hypoxia-driven partial EMT phenotype, we compared global gene expression profiles obtained from microarray platform of large hypoxic microtumors on early (day 1, 600/D1) and late time-points (day 6, 600/D6) with that of non-hypoxic small microtumors (150/D6). We identified enrichment of hypoxia, partial

EMT and migration gene signatures early on day 1 (600/D1) and at later time point (600/D6) in large microtumours. We identified *P4HA2*, a hypoxia-responsive gene as a central regulator that showed overlap between hypoxia, partial EMT and migration gene signatures. Further, inhibition of P4HA significantly abrogated migration in large microtumors supporting its role in partial EMT and collective migration.

## 2. Materials and methods

### 2.1. Cell culture

T47D cell line was purchased from American type culture collection (ATCC). Characterization and authentication of T47D cells was done at the University of Arizona Genomic Core facility using PowerPlex16HS PCR kit as previously described [22,23]. All cell culture supplies and media were procured from Corning® and Mediatech®, respectively, unless otherwise specified. T47D cells were passaged and maintained in T75 flask in Dulbecco's modified Eagle medium (DMEM) complete growth media (MT10013CV, Corning®, USA) supplemented with 10% fetal bovine serum (FBS) (S11250, Atlanta Biologicals, USA) and 1% penicillin-streptomycin (300-002-CI, Corning®, USA) in a humidified incubator with 5% CO<sub>2</sub> at 37 °C. Cells were maintained at 40–60% confluency before seeding into size-controlled hydrogel microwell devices.

### 2.2. Generation of size-controlled microtumors

The size-controlled T47D microtumors with diameters 100–150  $\mu$ m ('small' microtumors) and 450–600  $\mu$ m ('large' microtumors) were microfabricated using non-adhesive polyethylene glycol (PEG) hydrogel microwell devices as previously described [12,21–23]. Briefly, uniform size 150  $\mu$ m and 600  $\mu$ m hydrogel microwell devices (1  $\times$  1 cm<sup>2</sup>) were fabricated using polydimethyl siloxane (PDMS) stamps containing posts with 1:1 aspect ratio of height: diameter. Polyethylene glycol dimethacrylate (PEGDMA, 1000Da, Polysciences Inc., USA) solution (20% w/v) with 1% w/v photoinitiator (Irgacure-1959, Ciba AG CH-4002, Switzerland) was photo-crosslinked under the PDMS stamps to form size-controlled hydrogel microwells using OmniCure S2000 curing station (200W Lamp, 5W/cm<sup>2</sup>, EXFO, Canada). The fabricated hydrogel microwell devices were sterilized in 70% ethanol under UV light in laminar hood for 1h. Post-sterilization, devices were washed thrice with Dulbecco's phosphate buffered saline (DPBS) without calcium and magnesium (#21-031-CV, Corning™, USA). Subsequently, T47D cells (1.0 $\times$ 10<sup>6</sup> T47D cells/50 $\mu$ L/device) were seeded on to 1  $\times$  1 cm<sup>2</sup> hydrogel microwell devices. Excess cells remaining on top of the gels were washed slowly with DPBS. These cell-seeded microwell devices were then cultured in a humidified incubator with 5% CO<sub>2</sub> at 37 °C for six days with replacement of 50% media every day.

To study how tumor-intrinsic hypoxia drives partial EMT phenotype, large 600  $\mu$ m microtumors were cultured over 6-day period. The small non-hypoxic and non-migratory 150  $\mu$ m microtumors [22,23] served as controls.

### 2.3. Bioinformatic analysis of gene expression profiling data

To investigate the role of tumor-intrinsic hypoxia in driving phenotypic plasticity and collective migration, we have performed microarray analysis of 150/D6, 600/D1 and 600/D6 microtumors [25]. In the current study, we analyzed microarray data deposited in the Gene Expression Omnibus database at the National Center for Biotechnology Information (GSE166211) [25] to establish the link between hypoxia, partial EMT and tissue migration by comparing 600/D1 *versus* 150/D6 and 600/D6 *versus* 150/D6 comparison groups.

#### 2.3.1. Identification of differentially expressed genes

Microarray data deposited in the Gene Expression Omnibus database

at the National Center for Biotechnology Information (GSE166211) [25] was used for the analysis reported in this study. The microarray analysis for identification of differentially expressed genes in 600/D1 versus 150/D6 as well as 600/D6 versus 150/D6 was carried out as described previously [25]. The statistical analysis on each comparison was performed using Transcriptome Analysis Console (TAC 4.0, ThermoFisher Scientific, USA) using one-way ANOVA. A cut-off *p*-value of 0.01 was used for hierarchical clustering. Differentially expressed genes were considered significant using *p*-value  $\leq 0.05$ , false discovery rate (FDR)  $\leq 0.05$ , and fold change  $\pm 2$ .

### 2.3.2. Statistical enrichment of hypoxia, partial EMT and tissue migration signatures

The gene ontology enrichment of hallmark hypoxia in 600/D1 versus 150/D6 and 600/D6 versus 150/D6 groups was carried out using the pipeline as described in our previous work [25]. We used partial EMT gene signatures previously identified for head and neck cancer [26] as a reference gene signature set to assess the partial EMT signatures in 600/D1 and 600/D6 microtumors, Gene ontology database for GO: Tissue migration (GO:0090130) [27,28] was used to assess the tissue migratory signatures of the microtumors.

### 2.3.3. ssGSEA scoring of biological pathways

Hallmark signatures from MsigDB were used to quantify the activity of different biological pathways using the ssGSEA for scoring and quantification of the same [29]. The ssGSEA function [30] in the python package gseapy (<https://github.com/zqfang/gseapy>) was used for scoring of the biological pathways.

## 2.4. Multicolor flow cytometry staining

Small and large microtumors were cultured as described under section 2.2. At end of day 1, 3, and 6 of the culture, the microtumors were collected, washed with DBPS and disaggregated by incubating in Trypsin-EDTA for 5–10 min. For flow cytometry, three devices of 150  $\mu\text{m}$  microtumors and 1 device of 600  $\mu\text{m}$  microtumors were pooled to obtain required number of cells for each replicate and 3 such biological replicates for each group (150/D1, 150/D3, 150/D6, 600/D1, 600/D3, 600/D6) were used. The disaggregated cells were resuspended in PBS and processed for flow cytometric staining. Non-specific binding of fluorochrome-conjugated antibodies was minimized by pre-incubating pelleted cell suspensions for 5 min with neat decanted (56 °C, 30 min) mouse serum (5  $\mu\text{L}$ ) [31]. Prior to intracellular cytokeratin or intracellular E-Cadherin staining, cells were stained for surface E-Cadherin (Biolegend, Cat. No. 147315) and fixed with 2% methanol-free formaldehyde (Polysciences, Warrington, PA). Cells were then permeabilized with 0.1% saponin (Beckman Coulter) in DPBS with 0.5% human serum albumin (10 min at room temperature). The permeabilized cell pellets were incubated with 5  $\mu\text{L}$  of neat mouse serum for 5 min, centrifuged and decanted. The cell pellet was disaggregated and incubated with 2  $\mu\text{L}$  of anti-pan cytokeratin-FITC (Beckman Coulter, Cat. No. IM2356U, Clone J1B3). A separate tube containing cells unstained for surface E-Cadherin was incubated after permeabilization with E-Cadherin antibody- (R&D Cat. No. FAB18381P, clone DECMA-1, ECD fluorochrome) to determine total (surface and intracellular) E-Cadherin expression. Finally, cells were resuspended to a concentration of  $5 \times 10^6$  cells/mL and DAPI (Invitrogen Cat. No. D1306) was added to the cells at a concentration of 10  $\mu\text{g}/\text{mL}$ . Cells were acquired on Fortessa SORP flow cytometer (BD Biosciences, San Diego, CA). The instrument was calibrated using CS and T beads (BD Biosciences, Cat. No. 650621), and PMT voltages were adjusted to predetermined target channels using the seventh peak of 8-peak Rainbow Calibration Particles (Spherotech, Lake Forest, IL, Cat. No. RCP-30-5A) as a reference point. FITC (BD Biosciences, Cat. No. 349502 for cytokeratin expression) and single stained BD CompBeads anti-mouse Ig $\kappa$  (BD Biosciences Cat. No. 51-90-9001229, for E-Cadherin-CF594 clone DECMA1 expression), and

unstained cells (DAPI only) were used as spectral compensation standards [32].

## 2.5. Analysis of FCS files from flow cytometry

For each sample, data were exported into FCS files and organized by experiment into playlists using VenturiOne software (Applied Cytometry Systems, Dinnington, UK, V7.3). The gate for marker positive events was determined such that all negative controls (isotypes) represented less than 1% of clean events. Analytical data (event count, percent positive, MFI) were exported to CSV files and imported into Excel, where they were associated with sample ID and reagent.

## 2.6. The Cancer Genome Atlas (TCGA) analysis

We analyzed The Cancer Genome Atlas (TCGA) cohort of breast cancer patients in a subtype specific manner (Luminal A, Luminal B and Basal subtypes) to establish correlation of P4HA2 with epithelial and mesenchymal signatures. We scored each patient sample belonging to a given subtype separately by calculating the epithelial and mesenchymal scores through ssGSEA on the corresponding epithelial or mesenchymal gene lists as provided in the study by Tan et al. [33].

## 2.7. Treatment of microtumors with prolyl 4-hydroxylase (P4HA) inhibitor

To investigate the role of hypoxia-induced partial EMT on collective migration in large microtumors, we studied the effect of P4HA2 inhibition, a common hypoxia and partial EMT marker identified through our bioinformatic analysis. P4HA2 was inhibited by using 20  $\mu\text{mol}/\text{L}$  1,4-dihydrophenanthroline-4-one-3-carboxylic acid (1,4-DPCA) (SC-200758, Santa Cruz, USA). 1,4-DPCA, is a selective inhibitor of prolyl 4-hydroxylase, which potentially targets HIF-1 $\alpha$  upon hydroxylation of proline residue. Large 600  $\mu\text{m}$  microtumors were treated with 1,4-DPCA (20  $\mu\text{M}$ ) from day 1 over six-day period with replenishment with every change of media. To evaluate the effect of P4HA inhibition (P4HA2i) on collective migration of microtumors, number of migrating and non-migrating microtumors were counted on day 3 and day 6 (control versus treated microtumor devices). In addition, the control versus treated microtumors were also imaged at 4x and 10x magnification (Zeiss Primo-vert) on day 3 and day 6, and distance of migration (d) was calculated as described previously [23] for evaluating migration kinetics. Post-treatment, microtumors were collected on day 3 and day 6, respectively from control (untreated) and P4HA2i-treated devices for mRNA isolation [23,34].

## 2.8. Quantitative real time-polymerase chain reaction (qRT-PCR)

To validate the data obtained from microarray experiment, we validated the expression changes using qRT-PCR (QuantStudio 3, Applied Biosystems, USA). For each experiment, microtumors cultured in  $1 \times 1 \text{ cm}^2$  devices were harvested from PEG hydrogel devices (at least two devices for 600  $\mu\text{m}$  microtumors and four devices for 150  $\mu\text{m}$  microtumors) and stored in  $-80 \text{ }^\circ\text{C}$  in Trizol. RNA was isolated using GeneJET RNA purification kit (#K0732, Thermo Scientific, USA) as per manufacturer's protocol. DNase I (#EN0521, ThermoScientific, USA) treatment was used to remove genomic DNA from the RNA samples as per manufacturer's protocol in thermocycler (T100™, Bio-Rad Laboratories, Inc., USA). RNA concentration was quantitated using Nanodrop (#840274200™, Thermo Scientific, USA) and RNA quality was determined by absorbance ratio at 260/280 nm. Primers for GAPDH (housekeeping gene), P4HA1 and P4HA2, were procured from KiCq-Start™ primers (H\_GAPDH3.3, H\_P4HA2.1 and H\_P4HA2.2, Sigma-Aldrich, USA). mRNA expression levels of above-mentioned genes were estimated using iTaq™ Universal SYBR® Green One-step qRT-PCR kit (#1725151, Bio-Rad Laboratories Inc., USA) per manufacturer's

protocol. Each reaction was run in triplicates ( $n = 3$ ) with no template control (NTC) and no reverse-transcriptase control (RTC). The complete experiment was repeated in triplicates ( $n = 3$ ). Relative fold change expression for individual genes was calculated using  $2^{-\Delta\Delta Ct}$  method [35] and analyzed for fold change expression with respective 600/D6 untreated samples as controls.

## 2.9. Statistical analysis

Statistical analysis was done using GraphPad Prism (v9.0). Descriptive statistics and statistical comparisons (Student's t-test, 2-tailed) were performed, and data are represented as mean  $\pm$  SEM (standard error of mean) from three replicates. Unpaired t-test was used to compare two groups and  $p$ -value less than 0.05 was considered significant.  $P$ -values were Bonferroni-corrected for multiple comparisons.

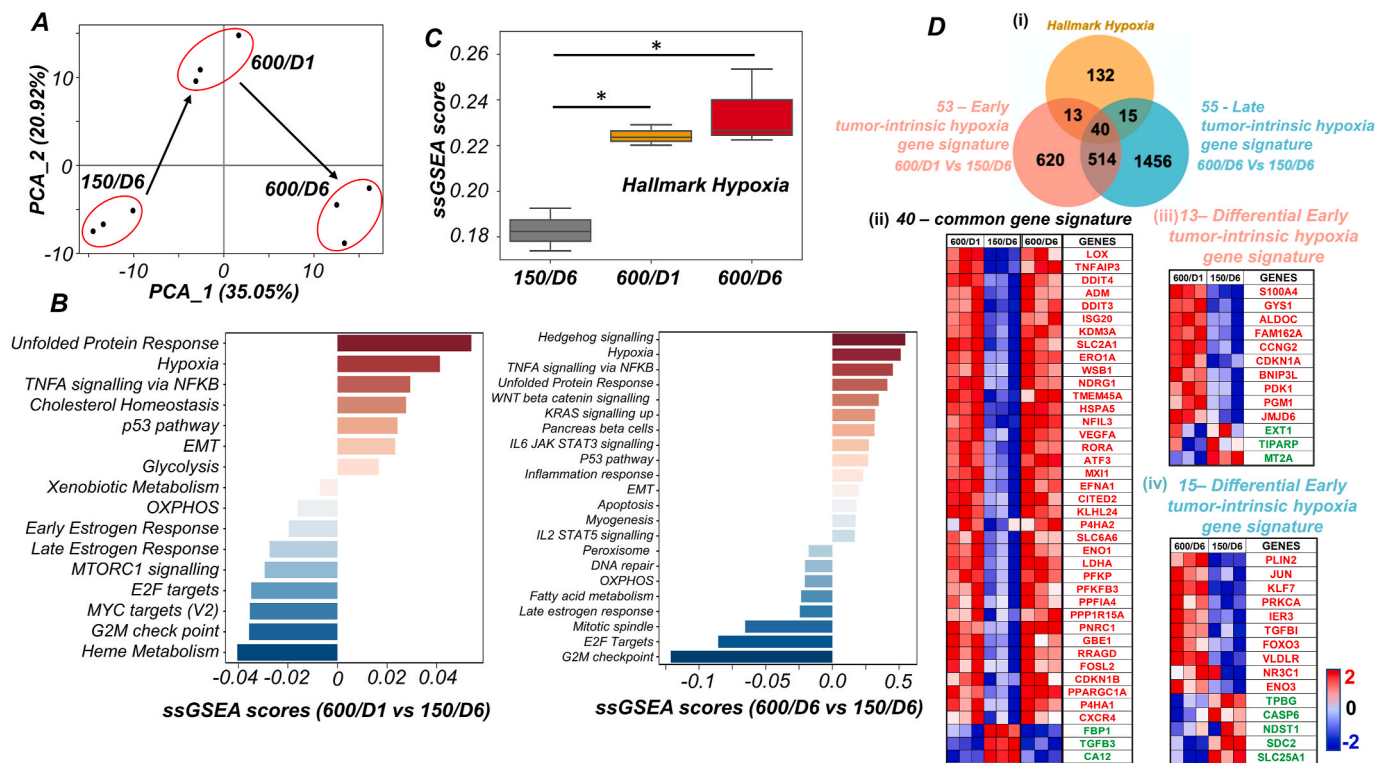
## 3. Results

### 3.1. Tumor-intrinsic hypoxia is enriched in large microtumors compared to small microtumors

To understand the transcriptomic differences between 150/D6, 600/D1 and 600/D6 microtumors, we performed Principal Component Analysis (PCA) on the gene expression matrices containing three biological replicates for each of the three microtumor models. Upon

plotting the samples on a two-dimensional PCA space, we found that  $\sim 56\%$  of the variance was explained by the first two PCA components (Fig. 1A). Furthermore, we found that the replicates clustered together on the PCA space with the inter-biological sample distances larger than the distances between replicates for a given biological sample. While 600/D6 microtumors varied from 150/D6 microtumors along the PCA-1 axis, the 600/D1 varied along both PCA-1 and PCA-2 with respect to either 150/D6 or 600/D6 samples. This difference indicates that some biological processes vary specifically in the 600/D1 and not in the 600/D6 samples, with 150/D6 samples as a control group. To investigate the biological processes that contribute to the variances in these biological samples, we calculated the ssGSEA scores for hallmark pathways for all the three microtumor models. Upon doing a differential analysis of scores between the pairs of microtumor groups 600/D1 versus 150/D6 and 600/D6 versus 150/D6 (Fig. 1B), we found that Hallmark Hypoxia was consistently upregulated in both these comparison sets, along with TNF $\alpha$  signaling, EMT, p53 signaling and unfolded protein response, although to varying degrees. This mapping shows that tumor-intrinsic hypoxia is upregulated as early as day 1 (600/D1 microtumors) and may contribute to initiate the migratory phenotype in these microtumors. To assess the extent of enrichment of hypoxia signature, we calculated the ssGSEA scores for the Hallmark Hypoxia and found that the hypoxia signature was significantly higher in the 600/D1 and 600/D6 compared to 150/D6 microtumors (Fig. 1C).

To obtain a comprehensive list of hypoxia gene signatures for large



**Fig. 1. Tumor-intrinsic hypoxia is highly enriched in large microtumors.** (A) Principal component analysis (PCA) plot for similarity assessment of gene expression levels across biological replicates of 150/D6, 600/D1 and 600/D6 microtumors fabricated from T47D cells; (B) ssGSEA enrichment of key biological processes in 600/D1 versus 150/D6 and 600/D6 versus 150/D6 comparison groups. Hallmark Hypoxia is enriched as early as day 1 and maintained over six-day period. Only hallmark gene sets, which show significant differences have been shown ( $p$ -value  $<$  0.05; Student's T-test); (C) Comparison of ssGSEA score analysis for Hallmark Hypoxia in 600/D1 and 600/D6 groups individually with 150/D6 microtumors as control shows significant enrichment of tumor-intrinsic hypoxia in both 600/D1 and 600/D6 groups. (D) (i) Venn diagram comparison of Hallmark Hypoxia with 600/D1 versus 150/D6 and 600/D6 versus 150/D6 comparison groups shows an overlap of 53 genes for 600/D1 (denoted as early tumor-intrinsic hypoxia gene signature) and 55 genes for 600/D6 (denoted as late tumor-intrinsic hypoxia gene signature). (ii) Forty genes are common among Hallmark Hypoxia, early and late tumor-intrinsic hypoxia gene signatures. (iii) Thirteen genes are differentially expressed only in 600/D1 versus 150/D6 and (iv) 15 genes are differentially expressed only in 600/D6 versus 150/D6. These are represented as heat maps with corresponding intensity values of genes from biological replicates in 600/D1 and 600/D6 versus 150/D6 group. In the heatmaps, 'red' represents upregulated status and 'blue' represents downregulated status. The corresponding gene names are illustrated in 'red' for upregulated status and 'green' for downregulated status. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

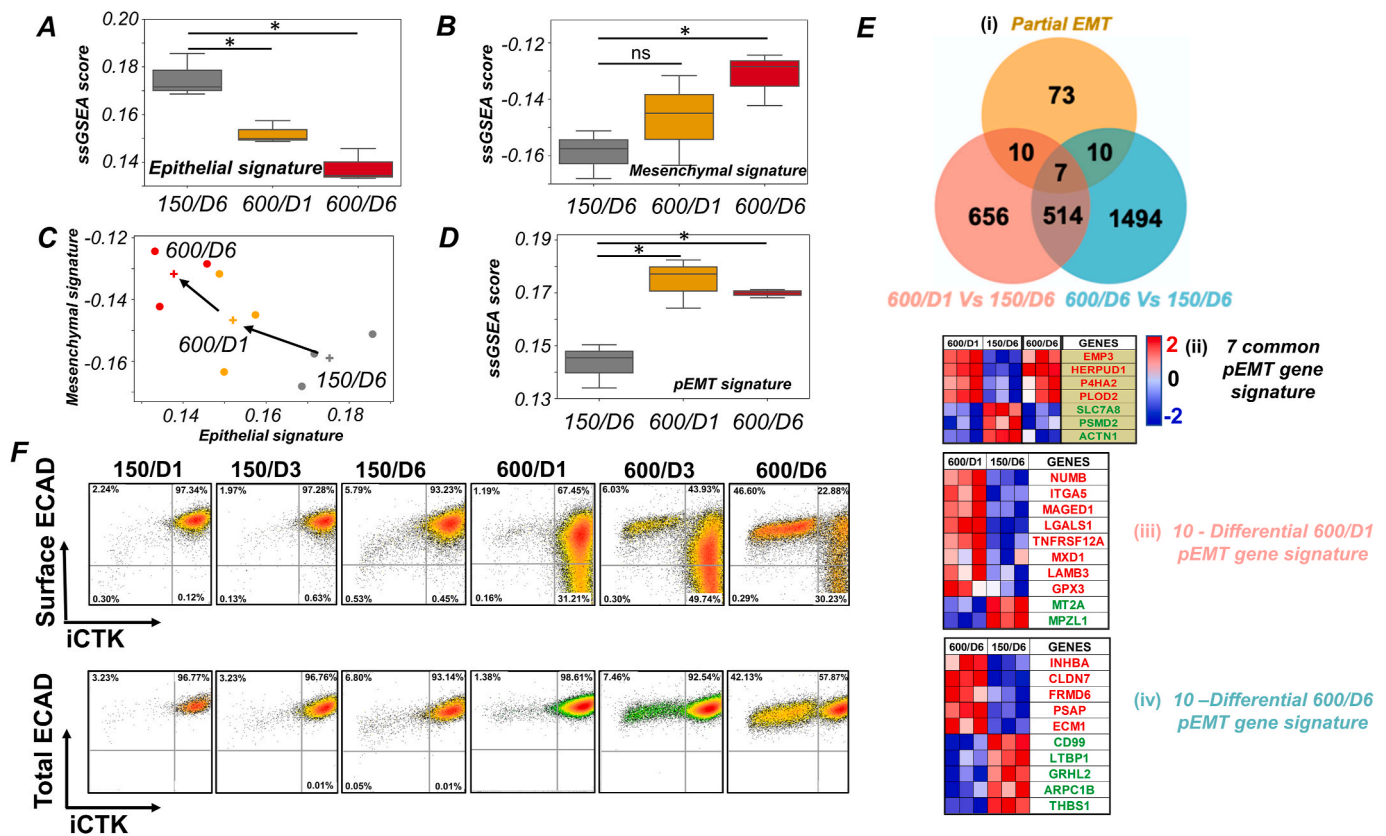
hypoxic microtumors (600/D1 and 600/D6) compared to small non-hypoxic microtumors (150/D6), we evaluated the overlap between Hallmark Hypoxia in 600/D1 versus 150/D6 and 600/D6 versus 150/D6 comparison groups (Fig. 1D). The gene signatures from Hallmark Hypoxia were collated to generate comprehensive hypoxia gene signature sets for 600/D1 versus 150/D6 (hereafter referred as “early tumor-intrinsic hypoxia gene signature”) and 600/D6 versus 150/D6 (hereafter referred to as “late tumor-intrinsic hypoxia gene signature”). A total of 53 genes were enriched in 600/D1 versus 150/D6 resulting in early tumor-intrinsic hypoxia gene signature while late tumor-intrinsic hypoxia gene signature was enriched with 55 genes in 600/D6 versus 150/D6 (Fig. 1D (i)). The Venn diagram-based comparative analysis identified 13 genes differentially expressed only in 600/D1 and 15 genes expressed only in 600/D6 (Fig. 1D (i)). We also found 40 genes common to both ‘early’ and ‘late’ tumor-intrinsic hypoxia gene signatures that are differentially expressed as early as day 1 (Fig. 1D (ii)). Of these 40 common genes, 37 genes were upregulated, and three genes were downregulated. Of the 13 genes differentially expressed only in 600/D1 (Fig. 1D (iii)), *S100A4*, *GYS1*, *ALDOC*, *FAM162A* and *CCNG2* were most upregulated while *MT2A*, *TIRAP* and *EXT1* were top downregulated genes (in descending order of expression). In the ‘late tumor-intrinsic

hypoxia’ signature of 600/D6 microtumors, *PLIN2* was the most upregulated gene followed by *JUN*, *KLF7*, *PRKCA* and *IER3* while *SLC25A1*, *SDC2*, *NDST1*, *CASP6* and *TPBG* were top five downregulated genes (in descending order of expression) (Fig. 1D (iv)). Taken together, gene expression analysis of large microtumor models (600/D1, 600/D6) with tumor size-induced natural hypoxia on day 1 and day 6 compared to non-hypoxic 150/D6 microtumors contributed ‘early’ and ‘late’ tumor-intrinsic hypoxia gene signatures.

### 3.2. Large microtumors exhibit enrichment of partial EMT signature in comparison to small microtumors

Our earlier work has demonstrated that the large 600  $\mu$ m microtumors with tumor-intrinsic hypoxia gain mesenchymal marker (VIM) without loss of epithelial marker (E-CAD), suggesting acquisition of partial EMT phenotype [23]. Based on these initial findings, we further investigated partial EMT signatures enriched by tumor-intrinsic hypoxia in large 600  $\mu$ m microtumors on day 1 and day 6.

First, we assessed changes in the epithelial and mesenchymal characteristics of 600/D1 and 600/D6 microtumors compared to control 150/D6 tumors. Specifically, we calculated the ssGSEA scores of

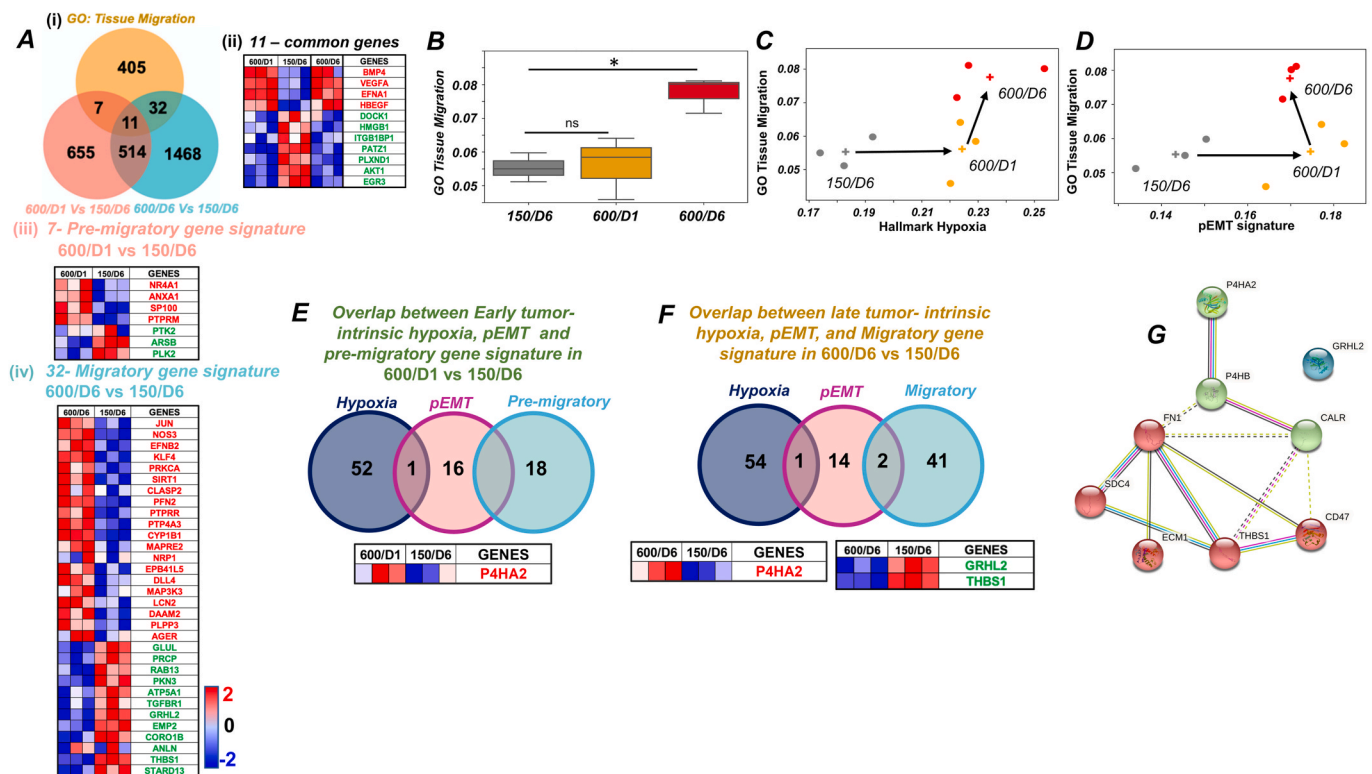


**Fig. 2. Large microtumors exhibit enrichment of partial EMT signature in comparison to small microtumors.** (A) ssGSEA score analysis for 'Epithelial signature' in 600/D1 and 600/D6 groups individually with 150/D6 microtumors as control shows significant decrease in epithelial signature over six-day period in both 600/D1 ( $p = 0.01^*$ ) and 600/D6 ( $p = 0.01^*$ ) microtumors; (B) Comparison of ssGSEA score analysis for 'Mesenchymal signature' in 600/D1 and 600/D6 groups individually with 150/D6 microtumors as control shows significant increase in expression of mesenchymal signatures over six-day period in 600/D6 ( $p = 0.01^*$ ) group but not in the 600/D1 group; (C) 2D plots showing Epithelial-Mesenchymal characteristics of 150/D6, 600/D1 and 600/D6 microtumors. (D) Comparison of ssGSEA score analysis for 'partial EMT signature' in 600/D1 and 600/D6 groups individually with 150/D6 microtumors as control shows significant increase in expression of partial EMT signatures [26] as early as day 1 in 600/D1 microtumors ( $p = 0.01^*$ ) and is maintained over six-day period in 600/D6 ( $p = 0.01^*$ ); (E) (i) Venn diagram comparison of partial EMT signature set [26] with 600/D1 versus 150/D6 and 600/D6 versus 150/D6 group shows that total 17 genes are enriched in 600/D1 versus 150/D6 as well as in 600/D6 versus 150/D6 group with an overlap of 7 genes in both these groups. Heatmap shows the intensity values of genes from biological replicates in 600/D1, 600/D6 and 150/D6 group as separate list for (ii) 7 genes common to both the groups, (iii) 10 differentially regulated genes present only in 600/D1 versus 150/D6 and (iv) 10 genes present only in 600/D6 versus 150/D6. In the heatmap, 'red' represents upregulated status and 'blue' represents down-regulated status. The corresponding gene names are illustrated with 'red' for upregulated status and 'green' for downregulated status; (F) Time-dependent expression of E-cadherin (E-CAD: surface ECAD and total ECAD) and intracellular pan-cytokeratin (iCTK) in T47D cells in 150  $\mu$ m and 600  $\mu$ m microtumors over day 1, 3 and 6. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

epithelial genes as reported by Tan et al. [33] and observed that there was a robust decline in the epithelial nature in 600/D1 and 600/D6 compared to 150/D6 tumors (Fig. 2A). Using the mesenchymal signatures from the same study [33], we found that the mesenchymal signature was upregulated significantly in 600/D6, but not in 600/D1 microtumors (Fig. 2B). This could indicate that the 600/D1 microtumors are more likely to be towards epithelial spectrum of partial EMT phenotype while the 600/D6 microtumors are more likely to be towards mesenchymal spectrum of partial EMT phenotype. To validate this, we plotted the samples on a two-dimensional plane of epithelial and mesenchymal phenotype scores. Indeed, 600/D1 tumors lie in an intermediate region between the more epithelial 150/D6 and the more mesenchymal 600/D6 tumors (Fig. 2C).

To identify partial EMT gene signatures in our breast microtumor models, we analyzed expression of partial EMT markers in 600/D1 and 600/D6 microtumors using available partial EMT gene signatures from head and neck cancer patients as a reference dataset [26]. We first calculated the enrichment of partial EMT signature using ssGSEA and found a significant increase in the partial EMT signature as early as

600/D1 compared to the 150/D6 tumors (Fig. 2D). We identified 17 partial EMT genes significantly enriched in 600 μm microtumors as early as day 1 and are referred to as ‘600/D1 partial EMT gene signature set’ (Fig. 2E). Over the six-day period in culture, non-migrating 600/D1 microtumors transitioned to collectively migrating phenotype (600/D6). Through comparison of gene expression in 600/D6 versus 150/D6, we identified enrichment of 17 partial EMT genes, referred to as ‘600/D6 partial EMT gene signature set’. The Venn diagram of comparative analysis of ‘600D1 partial EMT gene signature set’ and ‘600D6 partial EMT gene signature set’ identified 10 genes differentially expressed only in 600/D1 and 10 genes differentially expressed only in 600/D6 (Fig. 2E (i)). Seven partial EMT genes were common to both 600D1 and 600D6 microtumors, of which *EMP3*, *HERPUD1*, *P4HA2*, *PLOD2* were upregulated and *SLC7A8*, *PSMD2*, *ACTN1* were downregulated in both 600/D1 and 600/D6 microtumors compared to control 150/D6 microtumors (Fig. 2E (ii)). Of the 10 partial EMT genes differentially expressed only in 600D1 microtumors (Fig. 2E (iii)), *NUMB*, *ITGA5*, *MAGED1*, *LGALS1*, *TNFRSF12A*, *MXD1*, *LAMB3*, and *GPX3* were most upregulated while *MPZL1* and *MT2A* were downregulated (in



**Fig. 3. Tumor-intrinsic hypoxia drives partial EMT and collective migration in large microtumors.** (A) (i) Venn diagram comparison of GO: Tissue Migration with 600/D1 versus 150/D6 and 600D6 versus 150D6 comparison groups shows an overlap of 18 genes with 600/D1 (denoted as pre-migratory gene signature) and 43 genes with 600/D6 (denoted as migratory gene signature). (ii) Eleven genes are common among GO: Tissue Migration, pre-migratory and migratory gene signatures. (iii) Seven genes are differentially expressed only in 600/D1 versus 150/D6 and (iv) 32 genes are differentially expressed only in 600/D6 versus 150/D6. These are represented as heat maps with corresponding intensity values of genes from biological replicates in 600/D1 and 600/D6 versus 150/D6 group. In the heatmaps, ‘red’ represents upregulated status and ‘blue’ represents downregulated status. The corresponding gene names are illustrated in ‘red’ for upregulated status and ‘green’ for downregulated status. (B) Comparison of ssGSEA score analysis for GO Tissue Migration in 600/D1 and 600/D6 microtumors with 150/D6 microtumors as control shows significant increase in tissue migration only in 600/D6 ( $p = 0.005^*$ ) microtumors, but not in the 600/D1 microtumors; (C) 2D plot of Hallmark Hypoxia and GO: Tissue Migration showing the trajectory of transition of the microtumors on the phenotypic space of hypoxia and tissue migration; (D) 2D plot of partial EMT signature and GO:Tissue Migration showing the trajectory of transition of the microtumors on the phenotypic space of partial EMT and tissue migration; (E) Venn diagram comparison of early tumor-intrinsic hypoxia, partial EMT and pre-migratory gene signature set in 600/D1 versus 150/D6 shows an overlap of one gene signature (*P4HA2*) between tumor-intrinsic hypoxia and partial EMT. Heatmap shows the intensity values of genes from biological replicates in 600/D1 and 150/D6. The corresponding gene names are illustrated with ‘red’ for upregulated status. (F) Venn diagram comparison of late tumor-intrinsic hypoxia, partial EMT and migratory gene signature set in 600/D6 versus 150/D6 shows an overlap of one gene signature (*P4HA2*) between tumor-intrinsic hypoxia and partial EMT, and two genes (*GRHL2*, *THBS1*) between partial EMT and migratory gene signature. Heatmap shows the intensity values of genes from biological replicates in 600/D6 and 150/D6 group. In the heatmaps, ‘red’ represents upregulated status and ‘blue’ represents downregulated status. The corresponding gene names are illustrated with ‘red’ for upregulated status and ‘green’ for downregulated status. (G) Protein-protein interaction (PPI) network analysis of gene signatures common among hypoxia, partial EMT and tissue migration i.e. *P4HA2*, *GRHL2*, *THBS1*. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

descending order of expression) compared to 150/D6 microtumors. In '600/D6 partial EMT gene signature', *INHBA* was the most upregulated gene followed by *CLDN7*, *FRMD6*, *PSAP* and *ECM1* while *CD99*, *LTBP1*, *GRHL2*, *ARPC1B* and *THBS1* were top five downregulated genes (in descending order of expression) (Fig. 2E (iv)).

To further understand the expression profile of epithelial marker, E-CAD in our microtumor model, we evaluated surface and total (surface plus intracellular) expression of E-CAD in small (150  $\mu$ m) and large (600  $\mu$ m) microtumors. In small non-migratory microtumors, the expression profile of E-CAD did not change from day 1 to day 3 to day 6 (Fig. 2F, top panel). In large microtumors, we observed loss of surface E-CAD over six-day period, which is consistent with our prior work showing higher level of soluble (cleaved surface) E-CAD in the conditioned media of large microtumors [23]. However, analysis of combined surface plus intracellular E-CAD revealed no change in the total E-CAD (Fig. 2F, bottom panel).

### 3.3. Tumor-intrinsic hypoxia drives partial EMT and collective migration in large microtumors

To identify transcriptional signatures defining collective migration phenotype observed in the large hypoxic microtumors (600/D1, 600/D6) compared to small non-hypoxic microtumors (150/D6), we used the Gene Ontology signatures associated with tissue migration (GO: Tissue Migration) and evaluated the overlap between GO: Tissue Migration in 600/D1 versus 150/D6 and 600/D6 versus 150/D6 comparison groups (Fig. 3A). Compared to 150/D6 microtumors, we identified 18 gene differentially expressed in non-migratory 600/D1 microtumors (denoted as 'pre-migratory signature') and 43 genes differentially expressed in collectively migrating 600/D6 microtumors (denoted as 'migratory gene signature') (Fig. 3A (i)). Of these differentially expressed genes, 11 genes were commonly expressed in both the groups, which maintained their expression over six-day period (Fig. 3A (ii)). Of 18 genes, 7 genes were unique to non-migratory 600/D1 microtumors while out of 43 genes, 32 genes were unique to migratory 600/D6 microtumors (Fig. 3A(iii-iv)). We then calculated the ssGSEA scores for GO: Tissue Migration signatures and found significantly higher ssGSEA scores for the 600/D6 microtumors in comparison to 600/D1 microtumors (Fig. 3B). This was further validated in the two-dimensional plots of Tissue Migration with Hallmark Hypoxia signature as well as with the partial EMT signatures (Fig. 3C and D).

We further analyzed the overlap between hypoxia, partial EMT and tissue migration gene signatures in 600/D1 versus 150/D6 and 600/D6 versus 150/D6 groups to gain mechanistic insights into how tumor-intrinsic hypoxia contributes to partial EMT phenotype, which in turn, can modulate migratory phenotype. As shown in Fig. 3E, only *P4HA2* gene was common between early tumor-intrinsic hypoxia and partial EMT gene signature in 600/D1 microtumors. Indeed, *P4HA2* is known to be important for hypoxic adaptation and collagen biogenesis [36]. We did not find any overlap between partial EMT and pre-migratory gene signatures in 600/D1 microtumors (Fig. 3E). This analysis supports our observations, wherein 600/D1 microtumors exhibit tumor-intrinsic hypoxia and partial EMT; however, they are yet to migrate. As large microtumors start migrating by day 3, we observed enrichment of both partial EMT and migration gene signatures in 600/D6 microtumors. This may suggest that tumor-intrinsic hypoxia is responsible for initiating partial EMT and pre-migratory phenotypes, but acquisition of additional changes in the partial EMT genes may be required to further induce collective migration. Further evaluation of the overlap between late tumor-intrinsic hypoxia and 600/D6 partial EMT gene signatures revealed maintenance of *P4HA2* as a common gene shared between the two signatures (Fig. 3F). Thus, our analysis points to *P4HA2* as a central regulator of tumor-intrinsic hypoxia, partial EMT and migration. Few studies have highlighted role of *P4HA2* in driving EMT transition in glioma [37] and tumor migration in prostate cancer [38]; however, whether *P4HA2* drives migration through partial EMT states is not yet

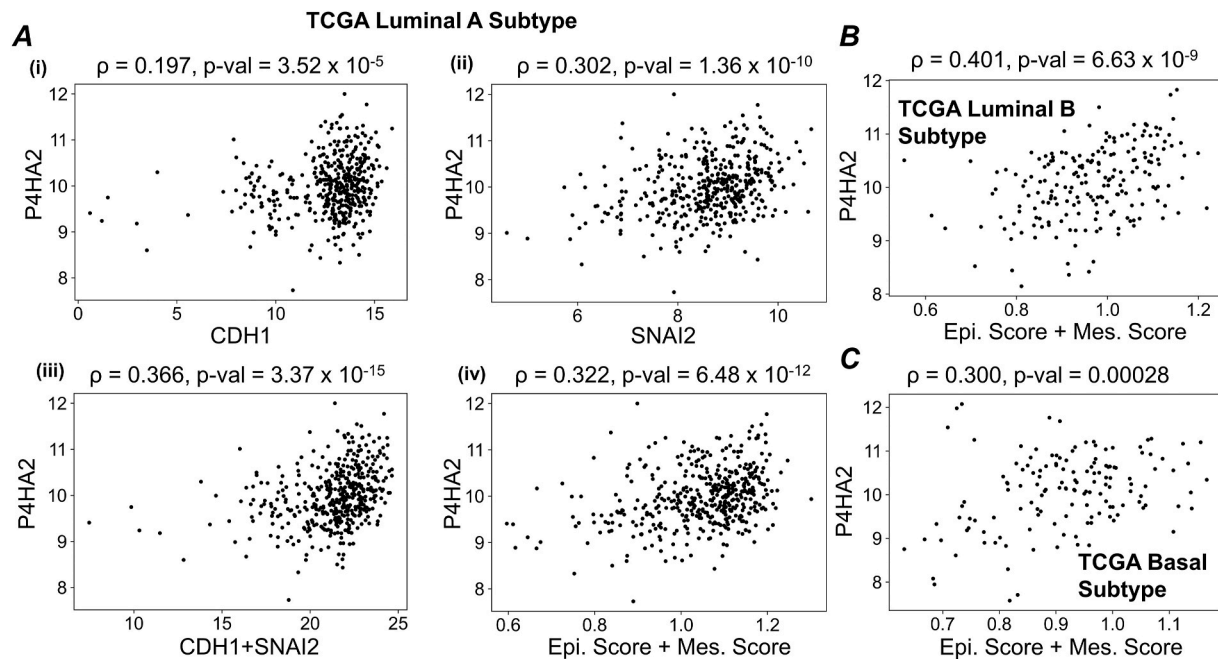
studied. Comprehensive gene expression analysis of hypoxic, yet non-migratory 600/D1 microtumors as well as that of hypoxic and migratory 600/D6 microtumors suggests that upregulation of *P4HA2* expression in response to tumor-intrinsic hypoxia plays a key role in inducing partial EMT states, which can further lead to collective migration. We also identified *GRHL2* and *THBS1* as common genes between partial EMT and collective migration phenotype in 600/D6 microtumors (Fig. 3F).

To further understand the protein-protein interaction (PPI) between these molecules, we performed *STRING* pathway analysis [39] (Fig. 3G). No known interaction was found between *P4HA2* and *GRHL2*; however, the PPI network highlighted direct experimental regulation of *P4HA2*, which exhibit protein-protein interactions with migration related genes namely, *THBS1*, *FN1* and *ECM1*, that were unique to 600/D6 migratory microtumors.

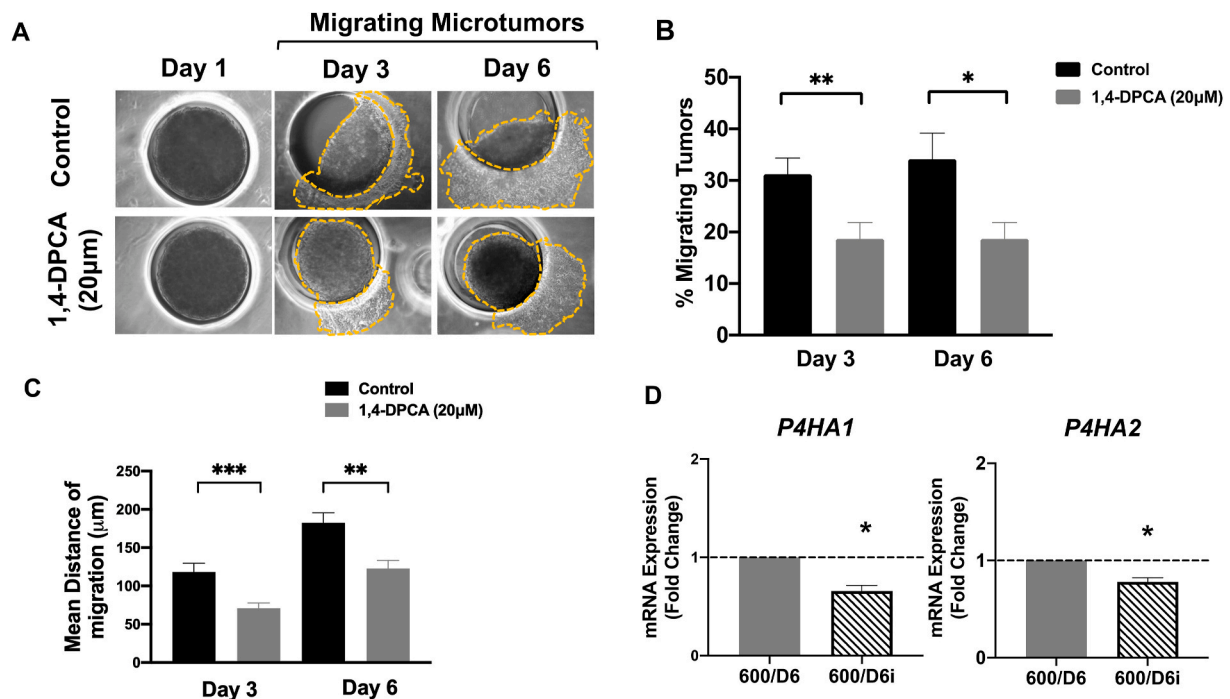
In addition, we also analyzed the TCGA breast cancer cohort of patients in a subtype specific manner (Luminal A, Luminal B and Basal subtypes) (Fig. 4). In Luminal A subtype of breast cancer, we show that *P4HA2* is correlated positively to both *CDH1* (an epithelial marker) and *SNAI2* (a mesenchymal marker) in patients (Fig. 4A (i-ii)). Furthermore, the sum of the expression of *SNAI2* and *CDH1*, which has been shown to be a characteristic of partial EMT state [40,41] revealed higher Spearman correlation coefficient with *P4HA2* expression than that for the individual correlations of *P4HA2* expression with *CDH1* and *SNAI2* (Fig. 4A (iii)). Finally, the sum of epithelial and mesenchymal ssGSEA scores (a proxy for partial EMT state – both scores being high) of the Luminal A subtype patients are also positively correlated with *P4HA2* expression (Fig. 4A (iv)). For Luminal B and Basal subtype of breast cancer, we see similar trends where *P4HA2* expression is correlated significantly positively with the sum of epithelial and mesenchymal ssGSEA scores (Fig. 4B and C).

### 3.4. Inhibition of *P4HA2* at early time point attenuates migration in large microtumors

Comparative analysis of common genes among early/late tumor-intrinsic hypoxia, partial EMT and migratory gene signatures (Fig. 3D and E) revealed that hypoxia-responsive *P4HA2* gene was upregulated in 600//D1 (Fig. 3D) and its expression was maintained over six-day period in 600//D6 (Fig. 3E). Therefore, we hypothesized that early inhibition of *P4HA2* will attenuate collective migration in the large microtumors. To study the effect of *P4HA2* inhibition on collective migration, we treated large microtumors with 1,4-DPCA (20  $\mu$ M), a P4HA inhibitor starting from day 1 to day 6. The photomicrographs of representative microtumors (Fig. 5A) illustrate comparative migratory phenotype in untreated (control) and 1,4-DPCA-treated 600  $\mu$ m microtumors in a time-dependent manner (D1, D3 and D6). In comparison to untreated microtumors, 1,4-DPCA-treated microtumors showed decrease in the distance of migration travelled by the microtumors from the edge of the well. The quantitative analysis of migrating microtumors post-1,4-DPCA treatment showed reduction in the number of migratory microtumors as early as day 3 ( $p = 0.0084^{**}$ ), which was sustained over 6 days ( $p = 0.011^{*}$ ) when compared to untreated microtumors (Fig. 5B). We also measured the distance of migration in the migratory microtumors treated with 1,4-DPCA in comparison to controls. The distance of migration was significantly reduced in 1,4-DPCA-treated microtumors compared to untreated controls on day 3 ( $p = 0.001^{***}$ ) and day 6 ( $p = 0.003^{**}$ ) (Fig. 5C). Quantitative RT-PCR studies showed significant decrease in the expression of both *P4HA1* and *P4HA2* in 1,4-DPCA treated microtumors compared to controls (Fig. 5D), indicating that 1,4-DPCA is a P4HA inhibitor that is not specific for *P4HA2*. Together, results from inhibition studies strongly suggest P4HA inhibition as an effective treatment strategy to block collective migration.



**Fig. 4.** Gene expression analysis of *P4HA2* with Epithelial and Mesenchymal genes in the TCGA cohort of breast cancer patients. (A (i-iv)) Scatterplots showing the correlations of *P4HA2* with *CDH1* (epithelial marker), *SNAI2* (mesenchymal marker), *CDH1+SNAI2* (a proxy for partial EMT phenotype) and sum of ssGSEA scores of Epithelial and Mesenchymal gene sets (a proxy for partial EMT phenotype) in the Luminal A subtype of breast cancer patients. Scatterplots showing the correlations of *P4HA2* with and sum of ssGSEA scores of Epithelial and Mesenchymal gene sets (a proxy for partial EMT phenotype) in the (B) Luminal B and (C) Basal subtype of breast cancer patients. Spearman correlation coefficients and the corresponding *p*-values are indicated for each scatterplot.



**Fig. 5.** Inhibition of *P4HA2* at early time point attenuates migration in large microtumors. (A) Photomicrographs of migrating 600  $\mu\text{m}$  microtumors in control (top panel) and 1,4-DPCA-treated (bottom panel) devices over day 1, day 3 and day 6; (B) Percentage migrating microtumors on 600/D3 and 600/D6 in control and 1,4-DPCA treated 600  $\mu\text{m}$  microtumors; (C) Effect of 1,4-DPCA on distance of migration travelled by 600  $\mu\text{m}$  microtumors evaluated on day 3 and day 6 for 10 individual tumors. 1,4-DPCA significantly reduced distance of migration on day 3 and day 6 as compared to the untreated microtumors; (D) Effect of 1,4-DPCA inhibition on the expression of *P4HA2* in treated 600  $\mu\text{m}$  microtumors. Fold change (mRNA expression) for *P4HA1* and *P4HA2* was determined by quantitative real-time PCR for microtumors harvested on day 6 in untreated and 1,4-DPCA treated 600/D6 microtumors. GAPDH, the housekeeping gene was used as a reference gene. Relative expression of *P4HA1* and *P4HA2* was compared to untreated 600/D6 controls (dotted line parallel to x-axis on bar graph). *P4HA1* and *P4HA2* expression decreased significantly in 1,4-DPCA treated 600  $\mu\text{m}$  microtumors. Data is represented as mean  $\pm$  SEM (Unpaired *t*-test, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001 compared to untreated 600/D6).



#### 4. Discussion

With updated guidelines in the field of EMT, a current focus is to understand the transiently stable partial EMT phenotypes wherein malignant cells acquire mesenchymal characteristics without a complete loss of epithelial markers [1]. Role of tumor microenvironmental factors such as tumor-intrinsic hypoxia, extracellular matrix or stromal factors in inducing or maintaining the partial EMT state remains an important understudied area of research [7]. Tumor-intrinsic hypoxia is one of the key contributors, if not the key driver and has long been known to drive tumor invasion and migration.

We have developed unique microtumor models that recapitulate tumor-intrinsic hypoxia using control over microtumor size. Small 150  $\mu\text{m}$  microtumors remain non-hypoxic and non-migratory over six-day culture period while large 600  $\mu\text{m}$  microtumors develop hypoxic signatures as early as day 1 [25] and exhibit collective migration starting from day 3 in culture [23]. We have also shown that once acquired, the migratory phenotype is irreversible [22] and that HIF1 inhibition at later time (after day 3) cannot block collective migration [23]. In this report, we exploited these unique microtumor models to investigate the link between tumor-intrinsic hypoxia, partial EMT and collective migration using integrated computational and experimental approaches.

From gene expression analysis, we propose a unique 53 gene-early tumor-intrinsic hypoxia gene signature and 55 gene-late tumor-intrinsic hypoxia gene signature. To the best of our knowledge, this is a first study to identify transcriptional changes in real time in response to naturally induced tumor-intrinsic hypoxia (in contrast to exposure to hypoxic chambers) at different time points as 3D microtumors transition from non-migratory to migratory phenotypes. We found 13 genes differentially expressed only in 600/D1 and 15 genes were expressed only in 600/D6. Of these 13 early tumor-intrinsic hypoxia gene signatures, elevated expression of *S100A4* has been correlated to hypoxia-induced tumor invasion and metastasis in oesophageal squamous cell carcinoma and ovarian cancer [42,43]. Other hypoxia-induced genes *GYS1* [44,45], *FAM162A* [46], *CCNG2* [47], *MT2A* [48,49] are known to be induced due to hypoxic response across different cancer types; however, their role as hypoxia responsive genes in breast cancer is not yet reported. Of the 15 late tumor-intrinsic hypoxia gene signatures, *PLIN2*, *JUN* and *PRKCA* are known to be induced in response to hypoxia [50–52].

We also identified transcriptional markers for partial EMT phenotype and propose a unique 17 gene-partial EMT gene signature for non-migratory 600/D1 and 17 gene-partial EMT gene signature for migratory 600/D6 microtumors. The ssGSEA analysis highlights phenotypic plasticity of T47D cells, which exhibit predominantly epithelial phenotype when cultured as non-hypoxic 150  $\mu\text{m}$  microtumors, epithelial-like partial EMT phenotype in hypoxic 600  $\mu\text{m}$  microtumors (day 1) transitioning to more mesenchymal-like partial EMT phenotype on day 6 as hypoxic 600  $\mu\text{m}$  microtumors become migratory. In large microtumors, we also observed loss of surface E-CAD over a six-day period, which is consistent with our prior work showing higher levels of soluble surface E-CAD in the conditioned media of large microtumors [23]. However, analysis of combined surface plus intracellular E-CAD revealed no change in the total E-CAD. Interestingly, similar observations have been reported in pancreatic ductal adenocarcinoma (PDAC) model, highlighting the loss of epithelial phenotype through E-CAD protein internalization instead of transcriptional repression of E-CAD and gain of mesenchymal markers resulting in a partial EMT phenotype [53]. The study further reported that the cells with partial EMT phenotype migrate as clusters, which closely resembles the findings from our 3D breast microtumor model, wherein large microtumors migrate as clusters and exhibit partial EMT phenotype. In a follow-up study using the same PDAC model system, calcium signaling – a key driver of the phenotypic stability factor (PSF), NFATc – was proposed to drive partial EMT [54].

Of note, we also observed differential regulation of some PSFs as

early as 600/D1 and also in 600/D6 microtumors. The PSFs stabilize or maintain the partial EMT phenotype in tumor cells and prevent complete EMT [55]. Recent studies have suggested the role of NUMB endocytic adaptor protein (*NUMB*), nuclear factor erythroid 2-related factor 2 (*NRF2*) and nuclear factor of activated T-cell (*NFATc*) as phenotypic stability factors [56,57]. Our microarray analysis of 600/D1 and 600/D6 microtumors revealed upregulation of *NUMB* [56] in hypoxic, but non-migratory 600/D1 microtumors supporting presence of partial EMT as early as day 1. In migratory microtumors (600/D6), *GRHL2* was found to be downregulated. While *GRHL2* was initially proposed as a PSF [8], more recent studies have indicated that it is more of an MET inducer than a PSF [58–60], thus its downregulation during later time is expected, as cells lose their epithelial traits progressively. In addition to *GRHL2*, we also identified *THBS1*, another PSF common between partial EMT and collective migration phenotype in 600/D6 microtumors. *THBS1* has been known to drive migration in cancer; however, its role as a partial EMT marker is not yet defined.

Our analysis of overlap between tumor-intrinsic hypoxia, partial EMT and migration signatures suggested *P4HA2* as a key link between tumor-intrinsic hypoxia, partial EMT and collective migration. Earlier studies have identified role of HIF-1 $\alpha$  in inducing *P4HA2* expression [61]. In glioma cells, *P4HA2* promoted EMT via collagen-dependent PI3K/AKT pathway and *P4HA2* knockdown inhibited migration, invasion and EMT [37]. However, whether *P4HA2* drives migration through partial EMT states is not yet studied. We identify *P4HA2* as a partial EMT marker induced by tumor-intrinsic hypoxia, and demonstrate its role in 3D collective migration through pharmacological inhibition studies. Thus, our study establishes a novel mechanism regulating hypoxia, partial EMT and collective migration.

In breast cancer patient microarray data, high mRNA expression of *P4HA2* was associated with collagen deposition (*COL1A1*, *COL3A1* and *COL4A1*) and poor patient outcomes [62]. Similar findings were reported in DCIS patients (pure versus mixed and invasive carcinoma) and *P4HA2* was suggested as an independent predictor of DCIS risk stratification [63]. These studies highlight the clinical relevance of *P4HA2* as a drug target in combination with established therapeutics.

#### 5. Conclusion

Taken together, using our unique 3D microtumor models, we show that tumor-intrinsic hypoxia changes transcriptional landscape driving partial EMT and collective migration. We identify *P4HA2*, known to be important for hypoxic adaptation and collagen biogenesis, as a key regulator of tumor-intrinsic hypoxia, partial EMT and collective migration, which may also serve as potential target to inhibit acquisition of migratory phenotype in breast cancer cells.

#### Funding

This work was supported by the National Institute of Health (NIH) [R37CA232209] to SS and VSD; the CDMRP BCRP BC210533, the MetaVivor foundation and the Glimmer of Hope foundation (VSD). MKJ was supported by Ramanujan Fellowship awarded by Science and Engineering Research Board (SERB), Department of Science & Technology (DST), Govt. of India (SB/S2/RJN-049/2018) and by InfoSys Foundation, Bangalore.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.



- sorting in immunological studies, *Eur. J. Immunol.* 47 (2017) 1584–1797, <https://doi.org/10.1002/eji.201646632>.
- [33] T.Z. Tan, Q.H. Miow, Y. Miki, T. Noda, S. Mori, R.Y. Huang, J.P. Thiery, Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients, *EMBO Mol. Med.* 6 (2014) 1279–1293, <https://doi.org/10.15252/emmm.201404208>.
- [34] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis, *Nat. Methods* 9 (2012) 676–682, <https://doi.org/10.1038/nmeth.2019>.
- [35] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, *Methods* 25 (2001) 402–408, <https://doi.org/10.1006/meth.2001.1262>.
- [36] J. Myllyharju, Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis, *Matrix Biol.* 22 (2003) 15–24, [https://doi.org/10.1016/s0945-053x\(03\)00006-4](https://doi.org/10.1016/s0945-053x(03)00006-4).
- [37] J. Lin, L. Jiang, X. Wang, W. Wei, C. Song, Y. Cui, X. Wu, G. Qiu, P4HA2 promotes epithelial-to-mesenchymal transition and glioma malignancy through the collagen-dependent PI3K/AKT pathway, *JAMA Oncol.* 2021 (2021), 1406853, <https://doi.org/10.1155/2021/1406853>.
- [38] M. Zhu, R. Peng, X. Liang, Z. Lan, M. Tang, P. Hou, J.H. Song, C.S.L. Mak, J. Park, S.E. Zheng, A. Huang, X. Ma, R. Chen, Q. Chang, C.J. Logothetis, A.K. Jain, S. H. Lin, H. Katayama, S. Hanash, G. Wang, P4HA2-induced prolyl hydroxylation suppresses YAP1-mediated prostate cancer cell migration, invasion, and metastasis, *Oncogene* 40 (2021) 6049–6056, <https://doi.org/10.1038/s41388-021-02000-3>.
- [39] D. Szklarczyk, A.L. Gable, K.C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, N. T. Doncheva, M. Legeay, T. Fang, P. Bork, L.J. Jensen, C. von Mering, The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets, *Nucleic Acids Res.* 49 (2021) D605–D612, <https://doi.org/10.1093/nar/gkaa1074>.
- [40] E. Sterneck, D.K. Poria, K. Balamurugan, Slug, E-Cadherin, Stealth accomplices? *Front. Mol. Biosci.* 7 (2020) 138, <https://doi.org/10.3389/fmolb.2020.00138>.
- [41] A.R. Subbalakshmi, S. Sahoo, K. Biswas, M.K. Jolly, A computational systems biology approach identifies SLUG as a mediator of partial epithelial-mesenchymal transition (EMT), *Cells Tissues Organs* (2021) 1–14, <https://doi.org/10.1159/000512520>.
- [42] A. Horiuchi, T. Hayashi, N. Kikuchi, A. Hayashi, C. Fuseya, T. Shiozawa, I. Konishi, Hypoxia upregulates ovarian cancer invasiveness via the binding of HIF-1 $\alpha$  to a hypoxia-induced, methylation-free hypoxia response element of S100A4 gene, *Int. J. Cancer* 131 (2012) 1755–1767, <https://doi.org/10.1002/ijc.27448>.
- [43] X. Xuan, Q. Li, Z. Zhang, Y. Du, P. Liu, Increased expression levels of S100A4 associated with hypoxia-induced invasion and metastasis in esophageal squamous cell cancer, *Tumour Biol* 35 (2014) 12535–12543, <https://doi.org/10.1007/s13277-014-2573-x>.
- [44] Y. Iida, K. Aoki, T. Asakura, K. Ueda, N. Yanai, S. Takakura, K. Yamada, A. Okamoto, T. Tanaka, K. Ohkawa, Hypoxia promotes glycogen synthesis and accumulation in human ovarian clear cell carcinoma, *Int. J. Oncol.* 40 (2012) 2122–2130, <https://doi.org/10.3892/ijo.2012.1406>.
- [45] N. Pescador, D. Villar, D. Cifuentes, M. Garcia-Rocha, A. Ortiz-Barahona, S. Vazquez, A. Ordonez, Y. Cuevas, D. Saez-Morales, M.L. Garcia-Bermejo, M. O. Landazuri, J. Guinovart, L. del Peso, Hypoxia promotes glycogen accumulation through hypoxia inducible factor (HIF)-mediated induction of glycogen synthase 1, *PLoS One* 5 (2010), e9644, <https://doi.org/10.1371/journal.pone.0009644>.
- [46] M.J. Lee, J.Y. Kim, K. Suk, J.H. Park, Identification of the hypoxia-inducible factor 1  $\alpha$ -responsive HGTD-P gene as a mediator in the mitochondrial apoptotic pathway, *Mol. Cell Biol.* 24 (2004) 3918–3927, <https://doi.org/10.1128/MCB.24.9.3918-3927.2004>.
- [47] A.Y. Ratushnyy, Y.V. Rudimova, L.B. Buravkova, Alteration of hypoxia-associated gene expression in replicatively senescent mesenchymal stromal cells under physiological oxygen level, *Biochemistry (Mosc.)* 84 (2019) 263–271, <https://doi.org/10.1134/S0006297919030088>.
- [48] A. Dube, J.F. Harrison, G. Saint-Gelais, C. Seguin, Hypoxia acts through multiple signaling pathways to induce metallothionein transactivation by the metal-responsive transcription factor-1 (MTF-1), *Biochem. Cell. Biol.* 89 (2011) 562–577, <https://doi.org/10.1139/o11-063>.
- [49] M. Yamasaki, T. Nomura, F. Sato, H. Mimata, Metallothionein is up-regulated under hypoxia and promotes the survival of human prostate cancer cells, *Oncol. Rep.* 18 (2007) 1145–1153, <https://doi.org/10.3892/or.18.5.1145>.
- [50] B. Qiu, D. Ackerman, D.J. Sanchez, B. Li, J.D. Ochocki, A. Grazioli, E. Bobrovnikova-Marjon, J.A. Diehl, B. Keith, M.C. Simon, HIF2 $\alpha$ -Dependent lipid storage promotes endoplasmic reticulum homeostasis in clear-cell renal cell carcinoma, *Cancer Discov.* 5 (2015) 652–667, <https://doi.org/10.1158/2159-8290.CD-14-1507>.
- [51] K. Sainikow, T. Kluz, M. Costa, D. Piquemal, Z.N. Demidenko, K. Xie, M. V. Blagosklonny, The regulation of hypoxic genes by calcium involves c-Jun/AP-1, which cooperates with hypoxia-inducible factor 1 in response to hypoxia, *Mol. Cell Biol.* 22 (2002) 1734–1741, <https://doi.org/10.1128/MCB.22.6.1734-1741.2002>.
- [52] F. Zeng, Y. Zhang, X. Han, M. Zeng, Y. Gao, J. Weng, Employing hypoxia characterization to predict tumour immune microenvironment, treatment sensitivity and prognosis in hepatocellular carcinoma, *Comput. Struct. Biotechnol. J.* 19 (2021) 2775–2789, <https://doi.org/10.1016/j.csbj.2021.03.033>.
- [53] N.M. Aiello, R. Maddipati, R.J. Norgard, D. Balli, J. Li, S. Yuan, T. Yamazoe, T. Black, A. Sahnoud, E.E. Furth, D. Bar-Sagi, B.Z. Stanger, EMT subtype influences epithelial plasticity and mode of cell migration, *Dev. Cell* 45 (2018) 681–695, <https://doi.org/10.1016/j.devcel.2018.05.027>, e4.
- [54] R.J. Norgard, J.R. Pitarresi, R. Maddipati, N.M. Aiello-Couzo, D. Balli, J. Li, T. Yamazoe, M.D. Wengyn, I.D. Millstein, I.W. Folkert, D.N. Rosario-Berrios, I. K. Kim, J.B. Bassett, R. Payne, C.T. Berry, X. Feng, K. Sun, M. Cioffi, P. Chakraborty, M.K. Jolly, J.S. Gutkind, D. Lyden, B.D. Freedman, J.K. Foskett, A. K. Rustgi, B.Z. Stanger, Calcium signaling induces a partial EMT, *EMBO Rep.* 22 (2021), e51872, <https://doi.org/10.15252/embr.202051872>.
- [55] P. Yaswen, Reinforcing targeted therapeutics with phenotypic stability factors, *Cell Cycle* 13 (2014) 3818–3822, <https://doi.org/10.4161/15384101.2014.985071>.
- [56] F. Bocci, M.K. Jolly, S.C. Tripathi, M. Aguilar, S.M. Hanash, H. Levine, J. N. Onuchic, Numb prevents a complete epithelial-mesenchymal transition by modulating Notch signalling, *J. R. Soc. Interface* 14 (2017), <https://doi.org/10.1098/rsif.2017.0512>.
- [57] A.R. Subbalakshmi, D. Kundnani, K. Biswas, A. Ghosh, S.M. Hanash, S.C. Tripathi, M.K. Jolly, NFATc acts as a non-canonical phenotypic stability factor for a hybrid epithelial/mesenchymal phenotype, *Front. Oncol.* 10 (2020), 553342, <https://doi.org/10.3389/fonc.2020.553342>.
- [58] V.Y. Chung, T.Z. Tan, M. Tan, M.K. Wong, K.T. Kuay, Z. Yang, J. Ye, J. Muller, C. M. Koh, E. Guccione, J.P. Thiery, R.Y. Huang, GRHL2-miR-200-ZEB1 maintains the epithelial status of ovarian cancer through transcriptional regulation and histone modification, *Sci. Rep.* 6 (2016), 19943, <https://doi.org/10.1038/srep19943>.
- [59] V.Y. Chung, T.Z. Tan, J. Ye, R.L. Huang, H.C. Lai, D. Kappei, H. Wollmann, E. Guccione, R.Y. Huang, The role of GRHL2 and epigenetic remodeling in epithelial-mesenchymal plasticity in ovarian cancer cells, *Commun Biol.* 2 (2019) 272, <https://doi.org/10.1038/s42003-019-0506-3>.
- [60] W. Jia, S. Tripathi, P. Chakraborty, A. Chedera, A. Rangarajan, H. Levine, M. K. Jolly, Epigenetic feedback and stochastic partitioning during cell division can drive resistance to EMT, *Oncotarget* 11 (2020) 2611–2624, <https://doi.org/10.18632/oncotarget.27651>.
- [61] H.Y. Jung, L. Fattet, J. Yang, Molecular pathways: linking tumor microenvironment to epithelial-mesenchymal transition in metastasis, *Clin. Cancer Res.* 21 (2015) 962–968, <https://doi.org/10.1158/1078-0432.CCR-13-3173>.
- [62] G. Xiong, L. Deng, J. Zhu, P.G. Rychahou, R. Xu, Prolyl-4-hydroxylase  $\alpha$  subunit 2 promotes breast cancer progression and metastasis by regulating collagen deposition, *BMC Cancer* 14 (1) (2014), <https://doi.org/10.1186/1471-2407-14-1>.
- [63] M.S. Toss, I.M. Miligy, K.L. Gorringe, A. Alkawaz, H. Khout, I.O. Ellis, A.R. Green, E.A. Rakha, Prolyl-4-hydroxylase  $\alpha$  subunit 2 (P4HA2) expression is a predictor of poor outcome in breast ductal carcinoma in situ (DCIS), *Br. J. Cancer* 119 (2018) 1518–1526, <https://doi.org/10.1038/s41416-018-0337-x>.