RaY

Research at the University of York St John

For more information please contact RaY at ray@yorksj.ac.uk

Pro- and anti-tumour activities of CD146/MCAM in breast cancer result from its heterogeneous expression and association with epithelial to mesenchymal transition

Aarren J. Mannion[1†], Adam F. Odell[1†], Syed Murtuza Baker[2], Laura C. Matthews[1], Pamela F. Jones[1‡] and Graham P. Cook[1*‡]

[1]Leeds Institute of Medical Research, University of Leeds School of Medicine, St. James's University Hospital, Leeds, United Kingdom, [2]Faculty of Biology, Medicine and Health, School of Biological Sciences, University of Manchester, Manchester, United Kingdom

CD146, also known as melanoma cell adhesion molecule (MCAM), is expressed in numerous cancers and has been implicated in the regulation of metastasis. We show that CD146 negatively regulates transendothelial migration (TEM) in breast cancer. This inhibitory activity is reflected by a reduction in MCAM gene expression and increased promoter methylation in tumour tissue compared to normal breast tissue. However, increased CD146/MCAM expression is associated with poor prognosis in breast cancer, a characteristic that is difficult to reconcile with inhibition of TEM by CD146 and its epigenetic silencing. Single cell transcriptome data revealed MCAM expression in multiple cell types, including the malignant cells, tumour vasculature and normal epithelium. MCAM expressing malignant cells were in the minority and expression was associated with epithelial to mesenchymal transition (EMT). Furthermore, gene expression signatures defining invasiveness and a stem cell-like phenotype were most strongly associated with mesenchymal-like tumour cells with low levels of MCAM mRNA, likely to represent a hybrid epithelial/mesenchymal (E/M) state. Our results show that high levels of MCAM gene expression are associated with poor prognosis in breast cancer because they reflect tumour vascularisation and high levels of EMT. We suggest that high levels of mesenchymal-like malignant cells reflect large populations of hybrid E/M cells and that low CD146 expression on these hybrid cells is permissive for TEM, aiding metastasis.

KEYWORDS

CD146, MCAM, transendothelial migration (TEM), epithelial mesenchymal transition (EMT), breast cancer, metastasis, cancer heterogeneity, tumour plasticity

Introduction

Metastatic disease is a hallmark of cancer and is responsible for the majority of cancer-related deaths (Hanahan and Weinberg, 2011; Dillekås et al., 2019). Metastasis occurs *via* the infiltration of malignant cells into surrounding tissue, their entry into the lymphatic or blood vessels (intravasation) and the dissemination to distant sites, where tumour cells exit these

vessels (extravasation) and seed the metastasis (Lambert et al., 2017). The crossing of endothelial barriers, termed transendothelial migration (TEM), involves the interaction between the migrating cell and endothelial cells (EC) and occurs in health during inflammatory responses (Vestweber, 2015). Several studies have shown that the TEM of cancer cells occurs *via* very similar mechanisms to those used by extravasating leucocytes (Madsen and Sahai, 2010; Reymond et al., 2013). TEM is a multi-step process mediated by a series of receptor-ligand interactions, cytoskeletal rearrangements and migratory activity, with active participation of both the migrating cell and the endothelium. These events result in migrating cells passing between (paracellular) and through (transcellular) EC to gain access to the tissues (Mamdouh et al., 2009; Reymond et al., 2013).

Numerous cell surface molecules expressed by both the migrating cell and EC are implicated in the regulation of TEM. For paracellular TEM, EC-EC interactions must be broken and both modes of TEM require interactions between the EC and migrating cell. The cell surface phenotype of tumour cells is thus a key factor in TEM and metastasis. Stable adhesion of cancer cells to the endothelium involves cancer cell surface molecules that are frequently over expressed in malignancy. For example, in breast cancer, MUC1 and CD44 overexpression facilitate tumour-EC interactions and promote TEM (Rahn et al., 2005; Wang et al., 2005; Yu et al., 2007; Zen et al., 2008), with MUC1 overexpression linked to poor prognosis (Jing et al., 2019). Along with multicomponent EC tight junctions and adherens junctions, EC-EC contacts are regulated by CD31, CD99 and CD146 and these molecules regulate TEM of inflammatory cells (Piali et al., 1995; Schenkel et al., 2002; Bardin et al., 2009). The CD146 molecule was first described as Melanoma Cell Adhesion Molecule (MCAM); this protein is highly upregulated in melanoma and was shown to mediate adhesion to EC (Xie et al., 1997). The CD146 molecule has subsequently been shown to have numerous functions in various cell types and, as such, plays a complex role in cancer progression (Wang and Yan, 2013).

An early step in metastasis is the generation of malignant cells with a migratory and invasive phenotype. For carcinomas, malignant epithelial cells can undergo epithelial to mesenchymal transition (EMT), differentiating them into mesenchymal-like cells which can detach from their epithelial neighbours and, having greater motility and invasive capacity, invade the surrounding tissue (Lamouille et al., 2014; Lambert and Weinberg, 2021). In addition, EMT promotes the acquisition of stem cell-like characteristics and drug resistance, generating cells with a potent capacity to seed metastases and resist treatment (Mani et al., 2008; Celià-Terrassa and Jolly, 2020). Not surprisingly, the expression of key genes regulating EMT are associated with patient outcomes (Moody et al., 2005; Taube et al., 2010; Lamouille et al., 2014; Lambert and Weinberg, 2021). Importantly, EMT is not characterised by a simple switch between epithelial and mesenchymal cells but is a spectrum of phenotypes from fully epithelial to fully mesenchymal. Indeed, stable intermediates can be identified, known as a hybrid epithelial/mesenchymal (E/M) or quasi-mesenchymal state and *in vivo* models have suggested that it is this hybrid state that has the strongest tumourigenic activity (Pastushenko et al., 2018; Kröger et al., 2019; Celià-Terrassa and Jolly, 2020; Lüönd et al., 2021; Pastushenko et al., 2021).

Furthermore, gene expression signatures characteristic of the hybrid state are markers of poor prognosis in breast cancer and several other solid tumours (Deshmukh et al., 2021; Zhang et al., 2022).

Here we have investigated the role of CD146 in the adhesion and TEM of breast cancer cells *in vitro*. Our results suggest that CD146 expression negatively regulates these events, a conclusion supported by reduced MCAM expression in breast cancer. However, patient-based survival data suggest a pro-tumorigenic role for CD146 in breast cancer. We demonstrate that these seemingly opposing roles for CD146 can be reconciled by considering the intra-tumoural heterogeneity of breast cancer and MCAM expression and the role of EMT in generating MCAM expressing, invasive cells.

Materials and methods

Cells and cell culture

Human umbilical vein ECs (HUVEC) were purchased from Promocell. Human cerebral microvascular endothelial cells (hCMEC/D3) referred to as hCMEC (VH Bio Ltd.) are an immortalised cell line isolated from human temporal lobe microvessels from tissue removed to treat epilepsy (Weksler et al., 2005). MDA-MB-231 cells were purchased from the European Collection of Cell Cultures and tested periodically for *Mycoplasma* contamination. Brain metastatic derived MDA-MB-231 cells (BrM), generated by serial *in vivo* passage of MDA-MB-231 cells in mice and subsequent isolation of cancer cells from metastatic lesions of the brain (Yoneda et al., 2001), were a kind gift from Dr. Mihaela Lorger (University of Leeds). EC lines were cultured using endothelial cell basal medium (ECBM; Promocell), supplemented with 2% foetal calf serum (FCS) (v/v), 0.4% Endothelial Cell Growth Supplement, 0.1 ng/mL epidermal growth factor (recombinant human), 1 ng/mL basic fibroblast growth factor (recombinant human), 90 μg/mL heparin and 1 μg/mL hydrocortisone. Cells were grown on 0.2% gelatin (Sigma-Aldrich) (w/v in PBS)-coated plates. HUVEC cells were grown to passage 5 or 6. hCMEC/D3 cell lines were grown to passage 35 before discarding as cells begin to lose endothelial characteristics (Weksler et al., 2005). MDA-MB-231 and BrM cells were cultured in 10% (v/v) FCS (Sigma-Aldrich)-supplemented RPMI-1640 (Sigma-Aldrich) and passaged every 3–5 days. All cell lines were incubated at 37°C under 5% CO_2.

Adhesion assays

Adhesion assays were carried out as previously described (Mannion et al., 2021). EC were seeded at a density of 10^4/well of a 96-well plate (Corning) and incubated until confluent monolayers were observed. MDA-MB-231 and BrM cells were labelled with 0.4 μM Cell Tracker Green (CTG) for 30 min in serum free RPMI (SFM-RPMI) medium at 37°C. MDA-MB-231 or BrM cells were washed in SFM RPMI once before being seeded at 10^4 per confluent EC monolayer. Adhesion assay was incubated at 37°C and CTG labelled cells were allowed to adhere to

the EC monolayers for 15, 30, 60 and 120 min, after which each plate was washed once in PBS, and fixed in 4% (w/v) paraformaldehyde (PFA; Sigma-Aldrich) for 10 min, and washed twice in PBS before storage at 4°C, followed by imaging using an Incucyte Zoom Live Cell Imager (Essen Bioscience). Images were subjected to ImageJ analysis, and the "watershed" function (www.imagej.net/Classic_Watershed) was used to distinguish between individual cells and clusters.

Transendothelial migration assay

24 well Thincert 3.0 μm or 5.0 μm pore diameter, transparent transwell filters (Greiner Bio-One Ltd.) were coated with 0.2% (w/v) gelatin and HUVEC or hCMEC/D3 were seeded at a density of 2×10^4 cells per insert. Cells were seeded in 300 μL ECBM media, with 500 μL in the lower chamber of the transwell insert. Endothelial cells were grown 24–48 h to allow formation of confluent monolayers before 2×10^4 breast cancer cells were seeded to the upper chamber. MDA-MB-231 and BrM cells (2×10^5/ml) were CTG labelled as described in (Mannion et al., 2021), before seeding to confluent EC monolayers in 1:1 ECMB:RPMI media. MDA-MB-231 and BrM cell migration was halted at 18 h by fixing in 4% PFA for 10 min followed by washing twice in 1x PBS (250 μL for upper chamber and 500 μL for lower chamber). Upper chambers of transwells were then scraped using cotton wool buds to remove cells on the upper layer of the transwell insert, leaving cells that had migrated to the underside of the membrane intact. Transwells were then washed twice in PBS. Migrated cells were then imaged using the EVOS microscope (Thermo Scientific).

Live cell imaging of intercalation

Cancer cell spreading and intercalation into endothelial monolayers is indicative of cancer cell transmigration (Reymond et al., 2012a; Onken et al., 2014; Mannion et al., 2021). Intercalation was determined by live cell imaging as previously described (Mannion et al., 2021; Mannion, 2022). Briefly, endothelial cells were seeded to 96 well plates at a density of 1×10^4/well in 100 μL to achieve confluent monolayers in 24–48 h. Once confluent endothelial monolayers were established, CTG labelled cancer cells (as described in adhesion assay) were seeded onto endothelial monolayers at a density of 1×10^4 per well in 50 μL of media (total media volume 150 μL including endothelial culture medium). Plates were then imaged immediately using Live Cell Imager—Incucyte Zoom. Images were taken every 5 min for 4 h using 20x objective. Images were analysed as previously described (Mannion et al., 2021).

RNA interference

MDA-MB-231 and BrM were transfected with SMARTpool siRNA (Dharmacon) targeting CD146 alongside a control scrambled (Scr) siRNA. Transfections were performed using Lipofectamine 200 RNAiMax (Invitrogen) transfection agent and Opti-MEM I Reduced Serum Medium, GlutaMAX Supplement

(Gibco) according to manufacturer's instructions. Cells were transfected with 30 pmol siRNA in a six-well plate ($2–4 \times 10^5$ cells/well) and scaled accordingly. Briefly, for a single well of a six-well plate, 30 pmol siRNA duplexes were made in 250 μL of OptiMEM medium and incubated at room temperature for 5 min. At the same time, 5 μL Lipofectamine was made up in 250 μL of OptiMEM and incubated at room temperature for 5 min siRNA and Lipofectamine mix were combined within the six-well plate and gently mixed before incubating at room temperature for 20 min. Following this, OptiMEM suspended cells were added to siRNA Lipofectamine complexes at $2–4 \times 10^5$ cells in 1 mL of OptiMEM. Cells were incubated in this mixture for 4–6 h, before transfection medium was aspirated and replaced with supplemented normal culture medium. siRNA-treated cells were incubated for 24–72 h before being used in downstream assays. The siRNA molecules used (from Dharmacon/Horizon Discovery) are shown in Supplementary Table S1.

Flow cytometry

Cultured cells were PBS washed and trypsinised with 1× Acutase (Gibco). Cells were washed in ice-cold PBS followed by centrifugation at $300\,g$ for 5 min. After repeated washing in PBS, cells were resuspended in 100 μL fluorescence-activated cell sorting buffer (PBS, 2% FCS and 0.09% NaN₃) and stained with fluorophore-conjugated antibodies (CD146-APC, SHM-57 or P1H12, BioLegend), (EPCAM-FITC, B29.1, VU-ID9, Abcam), (CD44-FITC, DB105, Miltenyi Biotec), (CD99-APC, HCD99 12E7, BioLegend) and relevant isotype control antibodies at 10^6 cells per 100 μL staining buffer for 30 min at room temperature. Stained cells were washed and fixed in Cytofix Fixation buffer (BD Biosciences) before analysis using a LSRII flow cytometer (BD Biosciences).

Patient sample-based gene expression, promoter methylation and survival analysis

Data from The Cancer Genome Atlas (TCGA) was accessed *via* The Cancer Immunome Atlas (TCIA) (Charoentong et al., 2017), available at https://tcia.at/home. In addition, we used RNAseq data from two other studies, one included multiple primary breast cancers, adjacent normal tissue and tissue from breast reduction surgery (Varley et al., 2014) and the other from a series of matched primary and brain metastases (Varešlija et al., 2019). For the latter, the normalised RNA-seq data provided by the authors was analysed directly. For the former, we downloaded metadata and raw short read archive (SRA) files (from Gene Expression Omnibus data series GSE58135), converted SRA files to FASTQ format and mapped them to human genome GRCh38 using STAR aligner v.2.5.1a. We used HTSeq v.0.10.0 to generate count matrices for genes across the samples. Raw counts were used for downstream data analysis in DESeq2; we created the DESeq2 object with raw counts with the cell metadata as the design matrix. We pre-filtered the reads that had at least 10 reads in total. To normalise, we used the median of ratio normalisation method within DESeq2 and applied variance stabilising transformation (VST) to stabilise variance across the mean. In addition, we validated tumour versus normal tissue

expression using the Gene Expression database of Normal and Tumour Tissue (GENT) 2 tool (Park et al., 2019), available at http://gent2.appex. kr/gent2/. MCAM promoter methylation was analysed using the Shiny Methylation Analysis Resource Tool (SMART) (Li et al., 2019), available at http://www.bioinfo-zs.com/smartapp/. For the association of gene expression with patient outcomes, we used the Kaplan Meier Plotter resource which incorporates breast cancer microarray data (Györffy et al., 2010) and breast cancer data from the pan-cancer RNA-seq data collection, both available at www.kmplot.com. For the TCGA cohort, clinical characteristics were obtained *via* the University of California, Santa Cruz Xena project (https://xenabrowser.net) and Kaplan-Meier survival analysis performed using GraphPad Prism. Gene expression was compared between patient groups using the statistical tests described in the figure legends, performed using GraphPad Prism software. For single cell (sc) analysis, we utilised scRNA-seq data from a study of 26 primary breast cancers (Wu et al., 2021). Data from this study was visualised, analysed and downloaded using Single Cell Portal (from the Broad Institute; https://singlecell.broadinstitute.org/single_cell). Single cell data was used to analyse expression of individual genes or to derive scores based on gene expression signatures. To sub-divide populations into MCAM^high, MCAM^low or MCAM^neg cells, we ranked MCAM expression within a group, marked cells without MCAM gene expression and divided the remaining MCAM expressing population into two equal size groups (for odd numbers of cells, we included an additional cell in the MCAM^low group).

Gene expression signatures

To determine the relative location of individual cells or tumours on the EMT spectrum we derived an EMT score (sEMT), calculated as the sum of expression of epithelial marker genes (CDH1, GRHL2, ITGB4, KRT5, KRT8, FST) subtracted from the sum of expression of mesenchymal marker genes (CDH2, ZEB1, VIM, MMP1, FN1, TGFB1I1). We used sEMT to classify malignant cells (or tumours) into sEMT^low, sEMT^med and sEMT^high sub-groups which likely represent populations enriched in epithelial-like cells, hybrid E/M cells and mesenchymal-like cells respectively. The precise boundaries used for this classification were set according to population distribution and are described in the relevant sections of Results. The utility of this method to derive sEMT was validated by analysis of other epithelial and mesenchymal genes as also described in Results. For the invasion score (sInv), we used the signature developed by Patsialou et al. (2012), with the exception that we only used genes shown to be upregulated in the invasive process. For the cancer stem cell score (sCSC), we used a twenty gene signature reported by Pece et al. (2019). Both sInv and sCSC were calculated as the mean expression of the genes in the respective signatures. An angiogenesis score (sAng), used by McDermott et al. (2018), was used to estimate tumour vasculature. A list of genes used to derive sEMT, sInv, sCSC and sAng are provided in Supplementary Table S2; the MMP1 gene appears in both sEMT and sCSC signatures, but otherwise the signatures are non-overlapping.

Statistical analysis

Statistical testing was performed using GraphPad Prism software and details of the parametric and non-parametric statistical tests used for different datasets are indicated in the figure legends. Data generated using the GENT2, SMART and KMplot databases was analysed using their own inbuilt statistical analysis tools (Györffy et al., 2010; Li et al., 2019; Park et al., 2019).

Results

We used a brain metastatic derivative of the triple negative breast cancer (TNBC) cell line MDA-MB-231 previously isolated from xenografts (Yoneda et al., 2001). This brain metastasis variant (here termed MDA-BrM) and the parental cell line (termed MDA) were analysed for their ability to adhere to human endothelial cell (EC) layers. Two types of EC were tested, one comprising umbilical vein endothelial cells (HUVEC) and a second using an immortalised EC line representing human cerebral microvascular endothelial cells (hCMEC/D3) orginally developed as a model of the endothelial component of the blood brain barrier (Weksler et al., 2005). The MDA and MDA-BrM cells were labelled with Cell Tracker Green (CTG) and seeded onto confluent HUVEC or hCMEC/D3 monolayers (Figure 1A) and the bound tumour cells quantified at various time points over a 2-h period (Figures 1B, C). Adhesion of both the MDA and MDA-BrM cells was greater on the HUVEC monolayers compared to the hCMEC/D3 monolayers as judged by the number of cells per field of view. Overall, the adhesion of MDA and MDA-BrM to either HUVEC or hCMEC/D3 was similar, although MDA showed significantly greater adhesion to HUVECs at the 60 min timepoint ($p < 0.05$), whereas MDA-BrM showed significantly greater adhesion to the hCMEC/D3 cells at 60 min ($p < 0.01$), reflecting the tropism of these tumour cells for particular tissues *in vivo*. However, these preferences were not evident after 120 min (Figures 1B, C).

We determined whether differential adhesion of MDA and MDA-BrM to EC monolayers impacted upon the ability of these tumour cells to undergo TEM using a transwell assay; HUVEC or hCMEC/D3 cells were grown to confluency on the upper membrane of the transwell chamber and CTG-labelled MDA or MDA-BrM added in serum free media. Lower chambers of the transwell contained 10% serum, providing a migratory stimulus to the tumour cells. Following an 18 h incubation, quantification of CTG-labelled tumour cells in the lower chamber revealed that MDA-BrM possessed significantly reduced TEM compared to the parental MDA line using both HUVEC ($p < 0.05$) and hCMEC/D3 ($p < 0.001$) endothelial barriers (Figure 1D). Importantly, no significant differences in serum-stimulated migration between MDA and MDA-BrM were identified in the absence of an EC barrier, indicating that differential TEM activity of MDA and MDA-BrM was due to interactions with the endothelial cell barrier rather than intrinsic differences in migratory activity (Figure 1D).

After initial adhesion, cells undergoing TEM exhibit a morphological change and spread over the endothelium. This is followed by migration between endothelial cells, a process termed intercalation; this can be followed *in vitro* using EC monolayers and live cell imaging (Reymond et al., 2012a; Mannion et al., 2021). CTG-labelled MDA and MDA-BrM cells were seeded at equal density onto confluent HUVEC monolayers and imaged over a period of 4 h to capture intercalation activity. Visual inspection of the images suggested that MDA-BrM was inferior at intercalation
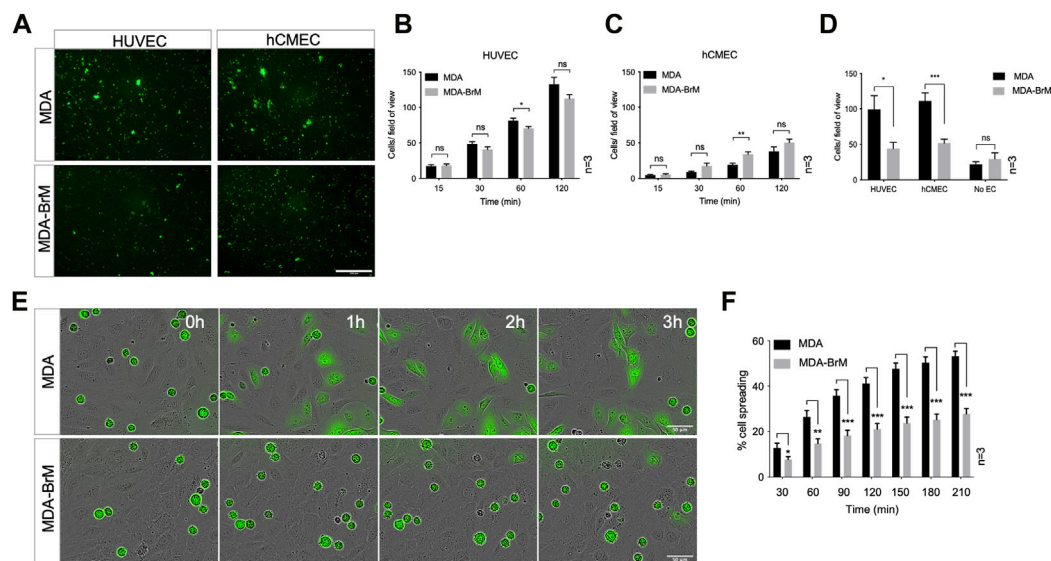
FIGURE 1
Adhesion of breast cancer cells to endothelial cells. **(A)** Adhesion of CTG labelled breast cancer cell lines MDA-MB-231 (MDA) and brain metastatic variant (MDA-BrM) to HUVEC and hCMEC endothelial monolayers. Images show adhesion at the 60 min time point. Scale bar: 500 μm. **(B)** Quantification of adhesion of MDA and MDA-BrM to HUVEC monolayers for indicated time points. Data are derived from three independent experiments. Error bars indicate standard error of the mean (S.E.M). Analysis was performed using multiple unpaired T testing; *p < 0.05; ns, not significant. **(C)** Quantification of adhesion of MDA and MDA-BrM to hCMEC monolayers for indicated time points. Data are derived from three independent experiments. Error bars indicate S.E.M. Analysis was performed using multiple unpaired T testing; **p < 0.01; ns, not significant. **(D)** Quantification of TEM of CTG labelled MDA or MDA-BrM seeded to HUVEC and hCMEC monolayers grown in the upper chamber of a Boyden transwell insert, or an empty transwell insert for No EC condition. CTG cells that had migrated to the underside of the transwell filter were imaged and quantified 18 h post-seeding. Data are derived from three independent experiments. Error bars indicate S.E.M. Analysis was performed using unpaired T testing; *p < 0.05; ***p < 0.001, ns, not significant. **(E)** MDA and MDA-BrM intercalation determined by live cell imaging. MDA and MDA-BrM were CTG labelled and seeded to HUVEC monolayers and intercalation/ spreading was captured using live cell imaging. Images were taken every 5 min for 3.5 h using a 20x objective. Scale bar: 50 μm. **(F)** Quantification of data in panel E, indicating the percentage of MDA or MDA-BrM cells that have undergone spreading/intercalation as a percentage of total cells. Data are derived from three independent experiments. Error bars indicate S.E.M. Analysis was performed using multiple unpaired T testing; *p < 0.05; **p < 0.01; ***p < 0.001.

into HUVEC monolayers (Figure 1E) and quantification confirmed that MDA-BrM cells were significantly impaired in intercalating activity in comparison to MDA at all time points analysed (p < 0.05–p < 0.001; Figure 1F). These results support the transwell migration assay data (Figure 1D), revealing that MDA-BrM cells have a greatly reduced capacity to undergo TEM in comparison to their parental counterparts.

A number of cell surface molecules have been implicated in the regulation of TEM, including CD44, CD99, CD155 and CD146 (Schenkel et al., 2002; Reymond et al., 2004; Wang et al., 2005; Zen et al., 2008; Bardin et al., 2009; Reymond et al., 2012b). We analysed the cell surface expression of these molecules, along with Ep-CAM, a marker of the metastatic phenotype and poor prognosis in breast cancer (van der Gun et al., 2010). Cell surface expression of CD44, CD99 and Ep-CAM was similar on MDA and MDA-BrM, but CD146 was expressed 8–10 fold higher on the MDA-BrM cells (p < 0.05; Figures 2A, B). This difference in expression was also seen at the total protein level (Figure 2C).

Cell surface CD146 participates in the TEM of inflammatory cells and melanoma cells and we speculated that it might also regulate the TEM of breast cancer cells. We used siRNA to inhibit CD146 expression and obtained a 75% reduction in cell surface CD146 in MDA-BrM cells (Figure 3A; p < 0.05). The MDA cells express a ~10 fold lower level of cell surface CD146 than MDA-

BrM and siRNA targeting reduced this expression by ~50% (Figure 3B; p < 0.05). We labelled the siRNA transfected MDA and MDA-BrM cells with CTG and performed a HUVEC adhesion assay; for both MDA and MDA-BrM, reduced cell surface expression of CD146 was associated with significantly increased adhesion to HUVEC monolayers at certain time points in the assay (p < 0.05–p < 0.0001; Figure 3C) revealing that CD146 expression inhibits breast cancer-EC adhesion. However, CD146 depletion did not alter intercalation into HUVEC monolayers for either cell line (Figure 3D).

We performed TEM assays in Boyden chambers using these siRNA treated cells; reduced expression of CD146 did not significantly affect TEM of MDA cells in this assay, (Figure 4A). However, for MDA-BrM, where unmanipulated CD146 expression was ~10 fold higher than in MDA, the reduction in CD146 expression resulted in a significant increase in TEM activity (Figure 4B; p < 0.01), a phenotype readily observed from the stained cell images.

However, when these Boyden chamber experiments were performed in the absence of EC, greater migratory activity was observed when CD146 expression was inhibited for both MDA and MDA-BrM, suggesting that the enhanced TEM of MDA-BrM resulting from CD146 knockdown was due to increased migratory activity rather than TEM itself (Figures 4A, B; p <
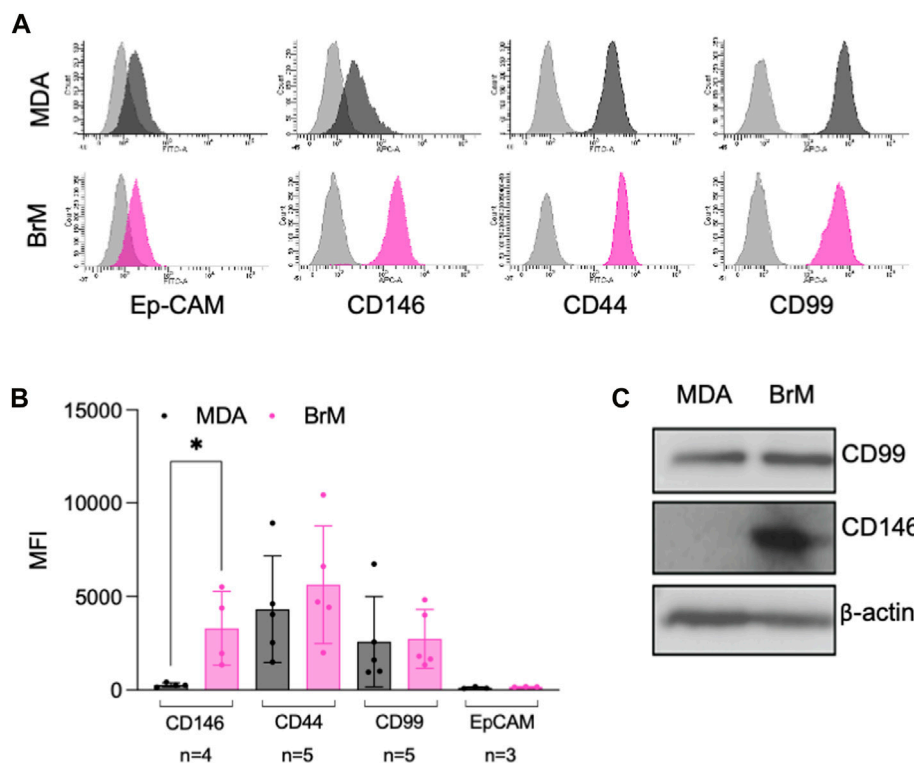
FIGURE 2
MCAM is highly expressed by brain metastatic variant of breast cancer **(A)** Expression of cell surface CD146, CD44, CD99 and Ep-CAM on MDA (dark grey histograms) and MDA-BrM (pink histograms) determined by flow cytometry compared to the isotype control (light grey histograms). This data is representative of a minimum of 3 independent experiments. **(B)** Quantification of flow cytometry data shown in **(A)**. The graph shows quantification of indicated receptor normalised to isotype controls with the number of independent experiments (n) used to measure expression of each antigen shown. Error bars indicate S.D. Analysis was performed using multiple unpaired T testing; *$p < 0.05$, ns, not significant. **(C)** Expression of total CD146 and CD99 protein expression determined by Western blotting in MDA and MDA-BrM using anti-CD146 and anti-CD99 antibody and anti-β actin as a loading control. This data is representative of two independent experiments.

0.05). These results show that CD146 expression inhibits the migration and TEM activity of MDA-BrM and suggests that CD146 expression functions as an inhibitor of these discrete stages of the metastatic process in breast cancer. Indeed, reduced CD146 expression allowed MDA-BrM cells to undergo TEM at a similar level to the parental MDA line, suggesting that the low levels of CD146 expressed by MDA cells are below the threshold of inhibition of TEM, whereas the high levels of CD146 on MDA-BrM are inhibitory.

To address the role of CD146 expression in breast cancer progression we analysed bulk tumour transcriptome data from patient samples. We analysed the expression of the MCAM gene (encoding CD146) across a panel comprising 42 oestrogen receptor (ER)+ primary breast cancer samples, 42 primary TNBC samples and 56 samples from normal adjacent tissue or non-cancerous breast tissue removed during breast reduction surgery (Varley et al., 2014). We first characterised the samples for expression of genes which define particular breast cancer types. By definition, TNBC lack expression of ER, the progesterone receptor (PR) and HER2; we analysed expression of the cognate genes (ESR1, PRG and ERBB2 respectively) in this dataset and found that all three genes were differentially expressed in the samples as expected. Furthermore, expression of the EPCAM gene, which is

overexpressed in breast cancer compared to normal tissue (van der Gun et al., 2010), was also differentially expressed (Supplementary Figure S1A). For the MCAM gene, we found differential expression across the three sample types (Figure 5A), with pairwise comparisons showing that MCAM expression was downregulated in both the ER+ ($p < 0.0001$) and TNBC samples ($p < 0.0001$) compared to the adjacent/normal breast tissue (Figure 5A). A significant reduction in MCAM gene expression in breast tumour compared to normal tissue was confirmed using two datasets from GENT2 (Park et al., 2019), a compendium of microarray data processed to allow comparisons between studies (Supplementary Figure S1B). In addition, we analysed RNA-seq data from matched pairs of primary breast cancer and their corresponding brain metastases (Varešlija et al., 2019). Breast cancer brain metastases upregulate KRT13 (Li et al., 2016) and downregulate CCDC8 (Pangeni et al., 2015), and these genes were differentially expressed in the primary and metastatic samples. However, MCAM expression was not significantly different between the primary tumour and the corresponding metastasis (Figure 5B).

Studies using breast cancer cell lines (including MDA-MB-231) have demonstrated that the MCAM gene is regulated by promoter methylation and that treatment with demethylating agents enhances MCAM gene expression and expression of CD146 (Dudzik et al.,
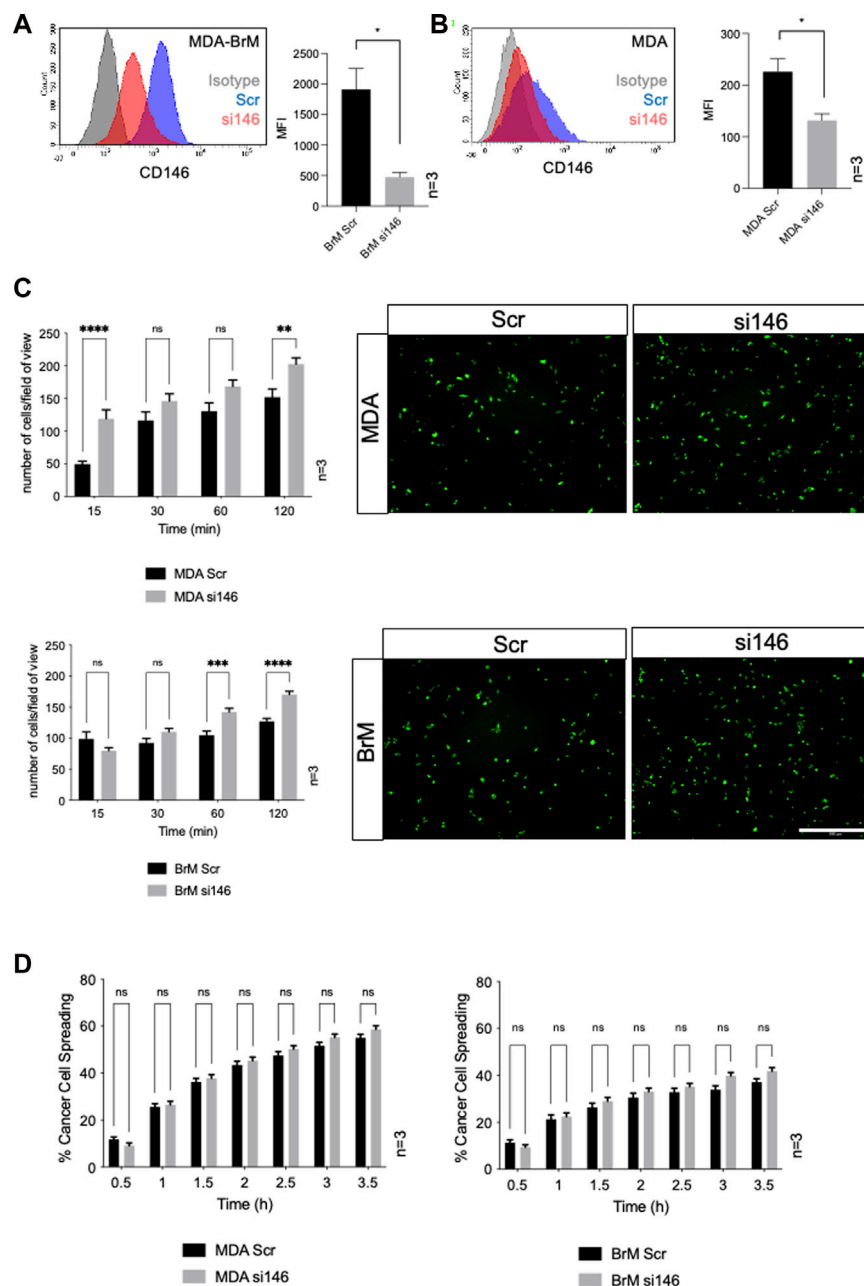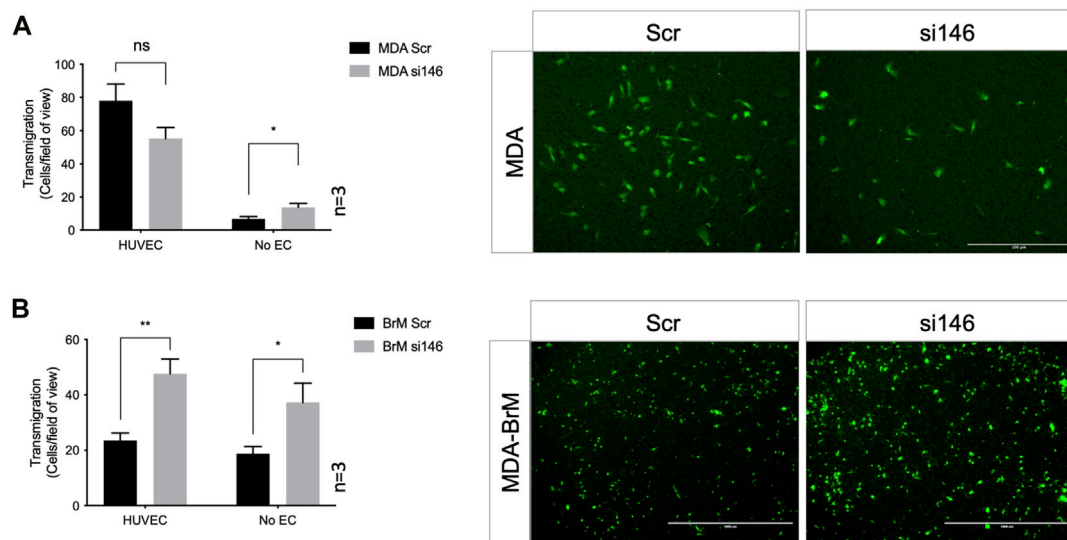
FIGURE 3
MCAM negatively regulates adhesion of breast cancer cells to endothelial cells **(A)** MDA-BrM cells were transiently transfected with siRNA targeting CD146 (si146) or a scrambled control siRNA (Scr), and CD146 expression was determined using flow cytometry 72−96 h post-transfection, compared to isotype control (grey histogram). For quantification, the mean fluorescence intensity (MFI) of CD146 expression in the Scr-treated cells was compared to si146-treated cells. Bar charts show data derived from three independent experiments. Error bars indicate S.D. Analysis was performed using an unpaired T-test; *$p < 0.05$. **(B)** Inhibition of CD146 expression in MDA cells, details as in **(A)**. **(C)** Adhesion of siRNA transfected MDA (top panels) or MDA-BrM cells (lower panels) transfected with CD146 (si146) or scrambled (Scr) control siRNA to HUVEC monolayers for indicated time points. The graphs show data from three separate experiments, error bars indicate S.E.M. Data were analysed using multiple unpaired T tests; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns, not significant. The images show adhesion of CTG labelled MDA or MDA-BrM transfected cells adhering to unlabelled HUVEC monolayers at the 60 min time point. Scale bar: 500 μm. **(D)** Quantification of MDA (left hand graph) and MDA-BrM (right hand graph) transfected with CD146 (si146) or scrambled (Scr) control siRNA and subsequent intercalation determined by live cell imaging. siRNA treated MDA and MDA-BrM were CTG labelled and seeded to HUVEC monolayers and intercalation/spreading was captured using live cell imaging. Quantification indicates the percentage of MDA or MDA-BrM cells that have undergone intercalation as a percentage of total cells. Images were taken every 5 min for 3.5 h using a 20x objective. Error bars indicate S.E.M. Data was derived from three independent experiments and analysed using multiple unpaired T tests; ns, not significant.

2019). This suggested that the reduced expression of the MCAM gene found in patient-derived breast cancer samples might be due to increased promoter methylation. We analysed methylation across a

CpG island spanning the transcriptional start site (TSS) of MCAM using the Shiny Methylation Analysis Resource Tool (SMART), which integrates methylation and expression data from The Cancer

FIGURE 4
MCAM negatively regulates TEM of brain metastatic variant of breast cancer **(A)** MDA cells were transiently transfected with siRNA targeting CD146 (si146) or a scrambled control siRNA (Scr), CTG labelled, and seeded to the upper chamber of a Boyden transwell chamber in the presence of HUVEC monolayers. For the No EC condition, MDA cells were seeded to empty transwells only. CTG cells that had migrated to the underside of the transwell filter were imaged (as shown) and quantified 18 h post-seeding, as shown in the bar diagrams. Scale bar: 200 μm. Data were derived from three independent experiments. Error bars indicate S.E.M. Analysis was performed using an unpaired T-test: *$p < 0.05$; ns, not significant. **(B)** As in **(A)** but using MDA-BrM cells. Scale bar: 1,000 μm. Data were derived from three independent experiments. Error bars indicate S.E.M. Analysis was performed using an unpaired T-test; *$p < 0.05$; **$p < 0.01$.

Genome Atlas (TCGA) (Li et al., 2019). We found significantly increased methylation in this region of the MCAM gene in cancer compared to normal tissue (Figure 5C; Supplementary Figure S2). Furthermore, MCAM gene expression was significantly and inversely correlated with methylation of this CpG island (Figure 5D).

The ability of cell surface CD146 to inhibit breast cancer TEM is consistent with the reduced expression of the MCAM gene in malignant versus normal breast tissue *via* epigenetic silencing. These results suggest that reduced CD146 expression in breast cancer might be a marker of poor prognosis. However, de Kruijff et al. (2018) reported the opposite, finding that high CD146 expression (as determined by immunohistochemistry) is associated with reduced overall survival and reduced metastasis free survival in breast cancer. We performed survival analysis based on MCAM gene expression and confirmed that high MCAM gene expression was associated with significantly reduced overall survival and distant metastasis free survival when combining multiple breast cancer types (Figures 5E, F; $p < 0.05$). For particular breast cancer subtypes (classified by gene expression in KMplot; 41), we found that high MCAM expression significantly reduced overall survival in HER2+ ($p < 0.01$) and TNBC ($p < 0.05$), but not ER + tumours (Supplementary Figure S3A) and that high MCAM gene expression was significantly associated with a poor outcome when analysed for distant metastasis free survival in HER2+ tumours ($p < 0.01$), but not TNBC or ER + PR + tumours (Supplementary Figure S3B). In addition, a separate dataset (using RNAseq instead of microarray data) confirmed the association of high MCAM expression with reduced overall survival (Supplementary Figure S3C). These

results mirror those of the immunohistochemistry study (De Kruijff et al., 2018) and show that high expression of MCAM is a marker of poor prognosis and is associated with metastasis in breast cancer.

Our expression data and TEM studies suggest an anti-tumour role for CD146, whereas prognostic studies indicate a pro-tumour role. This contradiction might be explained by intra-tumoural heterogeneity of CD146/MCAM expression, with different populations of CD146 expressing cells contributing differently to disease progression. We analysed MCAM gene expression at the single cell level, using sc-RNAseq data from ~100,000 cells derived from 26 breast cancer patients (Wu et al., 2021), a dataset that includes malignant cells (~24,000), as well as normal epithelium (~4,000 cells), cancer associated fibroblasts (CAF), immune cells, endothelial cells (EC) and perivascular cell (PVC) types. This dataset was viewed and analysed using Single Cell Portal at the Broad Institute. Expression of MCAM was detected in multiple cell types in breast cancer, including the malignant and normal epithelial cells, as well as other cell types, with high level expression found in EC and PVC (Figure 6A; Supplementary Figure S4). High MCAM expression in EC and PVC might account for the poor prognosis of patients with high MCAM gene expression levels, reflecting greater vascularisation of certain tumours. High expression of both the EC marker KDR/VEGFR2 and the PVC marker CSPG4 showed significant association with poor overall survival (Figure 6B; $p < 0.05$), suggesting that high levels of MCAM gene expression reflect greater vascularisation of tumours and associated poor prognosis. This was confirmed using bulk tumour samples (1,093 breast cancer patients from TCGA), where MCAM gene expression was shown to be positively correlated with an
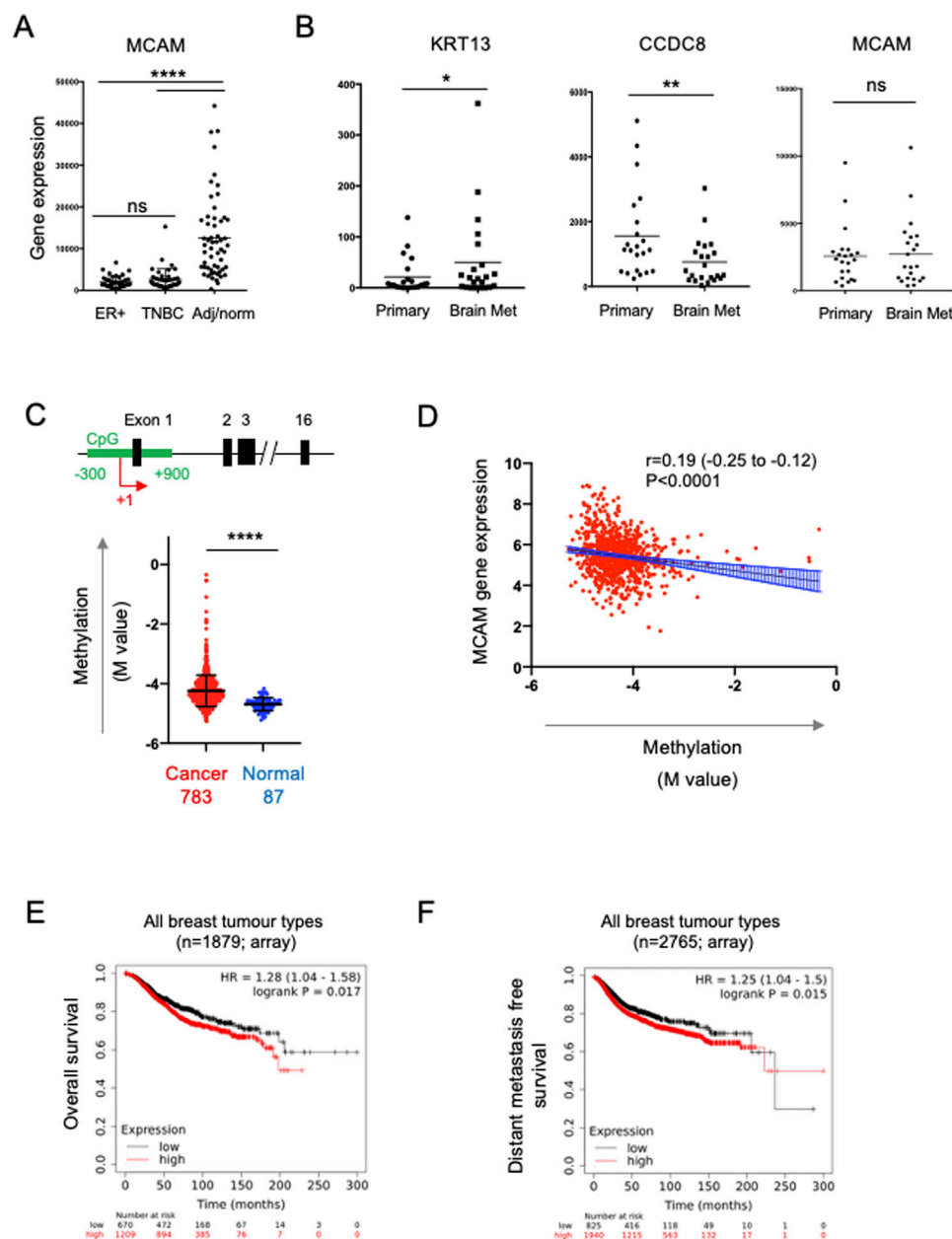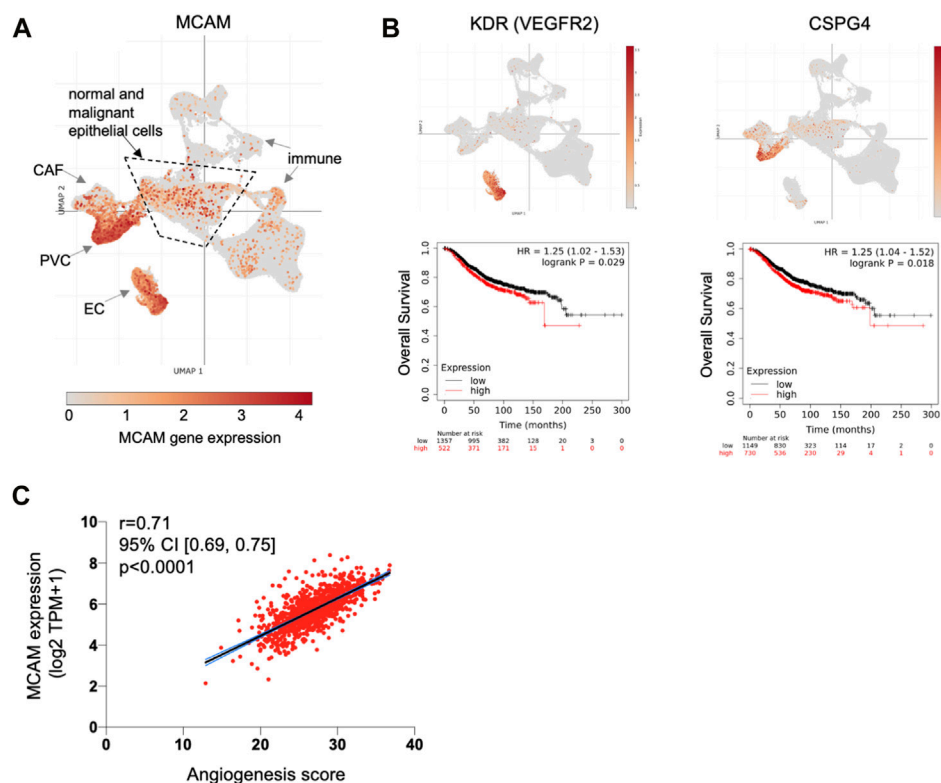
FIGURE 5
Expression of the MCAM gene in breast cancer. **(A)** MCAM gene expression in ER + breast tumours (*n* = 42), TNBC tumours (*n* = 42), as well as adjacent/normal breast tissue (*n* = 56). Expression values were obtained from the data of Varley et al. (2014). Differences in expression between the pairs of samples indicated was analysed using a two-tailed Mann-Whitney test; ****p < 0.0001 or not significant (ns). Further analysis of these samples for ESR1, PGR, ERBB2 and EPCAM gene expression is shown in Supplementary Figure S1. **(B)** Expression of KRT13, CCDC8 and MCAM genes in 26 primary breast cancers and patient matched brain metastases; expression values were obtained from the data of Vareŝlija et al. (2019). Differences in expression between these samples was analysed using a two-tailed Wilcoxon matched-pairs signed rank test; *p < 0.05, **p < 0.01 or not significant (ns). **(C)** Methylation of a CpG island around the transcriptional start site (TSS; +1) of the MCAM gene. The diagram shows a representation of the MCAM gene, indicating the position of the CpG island (in green) relative to the exons (black) and the TSS (+1, red arrow). The graph shows the methylation (M value) in breast cancer and normal tissue samples (793 and 87 samples respectively, from TCGA) using data from nine probes across the CpG island. For each sample, we calculated the mean M value and compared these values in the tumour and normal tissue using a two-tailed Mann Whitney test; ****p < 0.0001. Data analysis and download was performed using SMART (Li et al., 2019). Methylation data for individual probes within the CpG island are shown in Supplementary Figure S2. **(D)** Correlation of MCAM methylation with MCAM gene expression (using TCGA data analysed *via* SMART). Spearman's r, 95% confidence internals (and associated *p*-value) are shown together with the linear regression line and confidence intervals. **(E)** Kaplan-Meier analysis of overall survival in a cohort of 1879 breast cancer patients stratified for MCAM gene expression. **(F)** Kaplan-Meier analysis of distant metastasis free survival in a cohort of 2,765 patients containing all breast cancer types, stratified for MCAM gene expression. For **(E,F)**, data was graphed and analysed and using KM plotter (Győrffy et al., 2010).

FIGURE 6
MCAM gene expression at the single cell level in human breast cancer. **(A)** Uniform Manifold Approximation and Projection (UMAP) of a complete breast cancer scRNA-seq dataset comprising 100,064 cells from 26 patients. This data is from the study of Wu et al. (2021) with major cell populations indicated, as previously defined; these include immune cell types, cancer associated fibroblasts (CAF), endothelial cells (EC) and perivascular cells (PVC), as well as the epithelial cells (both malignant and normal) within the central area as indicated. Expression of MCAM is indicated and superimposed in orange. Data was displayed and analysed using Single Cell Portal. **(B)** UMAP of the complete sc-RNAseq dataset showing expression of the endothelial cell marker gene KDR (VEGFR2) and the PVC marker gene CSPG4 (superimposed in orange). Below each UMAP are Kaplan Meier plots showing overall survival in 1879 breast cancer patients stratified for KDR and CSPG4 gene expression, as determined using KM plotter (Győrffy et al., 2010). **(C)** Correlation of angiogenesis score with MCAM gene expression across a cohort of 1,093 breast cancer patients from TCGA. Spearman's r and the 95% confidence intervals (with associated *p*-value) are shown.

angiogenesis score (Figure 6C; Spearman's r = 0.71; 95% CI; 0.69–0.75, *p* < 0.0001).

Our *in vitro* TEM data shows that CD146 regulates the adhesion and migration properties of the tumour cells themselves. Furthermore, high levels of tumour cell CD146 are a marker of poor outcome in breast cancer (De Kruijff et al., 2018). We analysed MCAM gene expression within the epithelial cell populations in detail (using the sc-RNAseq data) and found that they were highly heterogeneous for MCAM expression; a greater proportion of normal epithelial cells (10%) expressed MCAM transcripts at detectable levels compared to their malignant counterparts (4%). Furthermore, the expression level of MCAM was reduced in the malignant epithelial cells compared to their normal counterparts, whereas for EPCAM, the opposite relationship was found (Figure 7A). We repeated the analysis of MCAM using the sc-RNAseq data from four individual patients included in the study, choosing samples where the number of malignant cells and normal epithelial cells both exceeded one hundred. These results confirmed that MCAM expression was significantly reduced in TNBC, ER+, and ER+/HER2+ breast cancer compared to the associated normal epithelium and demonstrated intra-tumoural heterogeneity in MCAM gene expression in both normal and malignant epithelial populations (Figure 7B).

One important driver of intra-tumoural heterogeneity is EMT (Mani et al., 2008; Lamouille et al., 2014; Celià-Terrassa and Jolly, 2020; Lambert and Weinberg, 2021). This is a dynamic, reversible process and, within a tumour, malignant cells occupy a variety of states across the EMT spectrum rather than simply being either epithelial or mesenchymal (Pastushenko et al., 2018; Kröger et al., 2019). We determined the relative position of each of the ~25,000 malignant cells across the EMT spectrum by deriving an EMT score (sEMT) for each cell based on the expression of twelve genes, six defining the epithelial phenotype and six from the mesenchymal phenotype. This analysis demonstrated that the malignant cells did indeed encompass a spectrum of sEMT and suggested the presence of three overlapping populations; sEMT^{low} (78% of cells; sEMT<0), sEMT^{med} (18%; sEMT 0–4.99) and sEMT^{high} (4%; sEMT>5), which likely represent populations enriched in epithelial-like cells, intermediate cells (hybrid E/M) and mesenchymal-like cells respectively (Figure 7C). We analysed these three populations defined by sEMT for the expression of
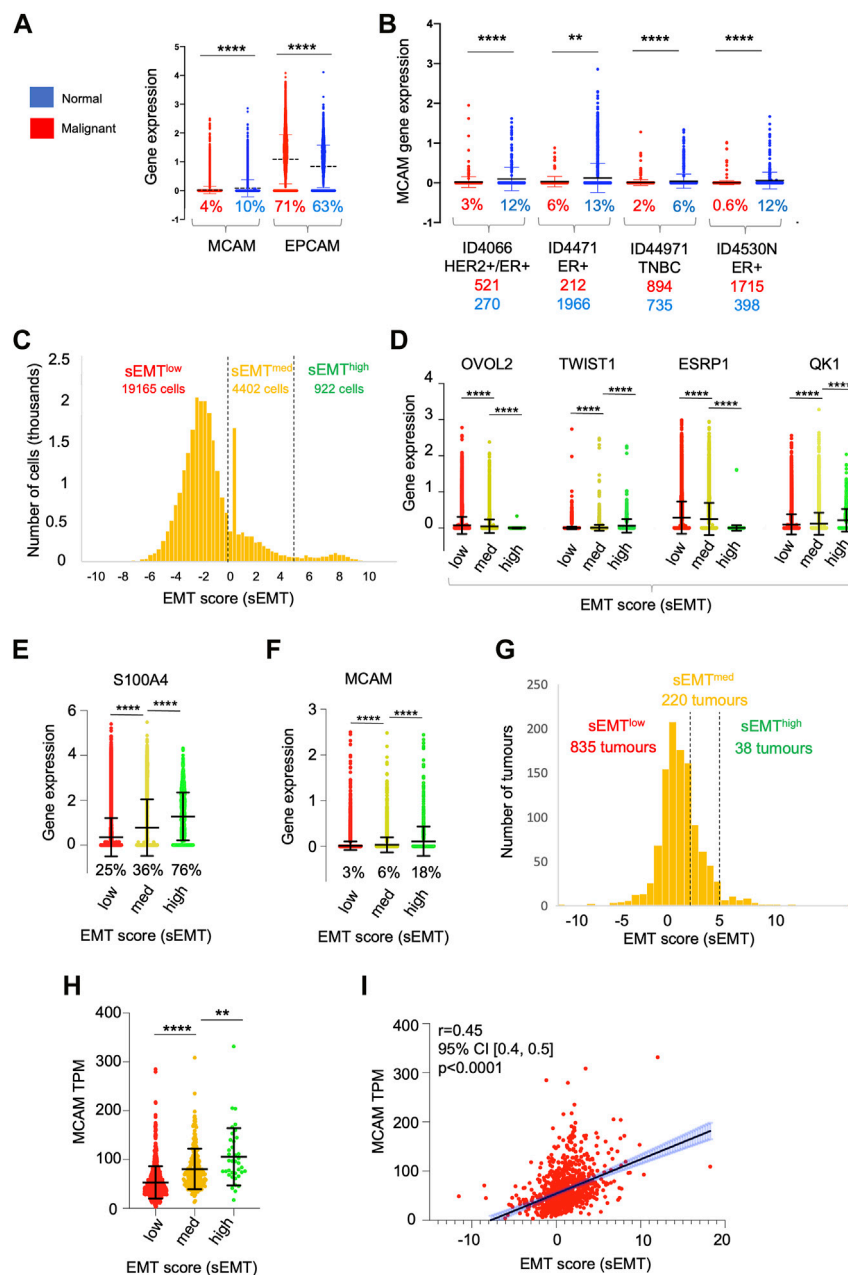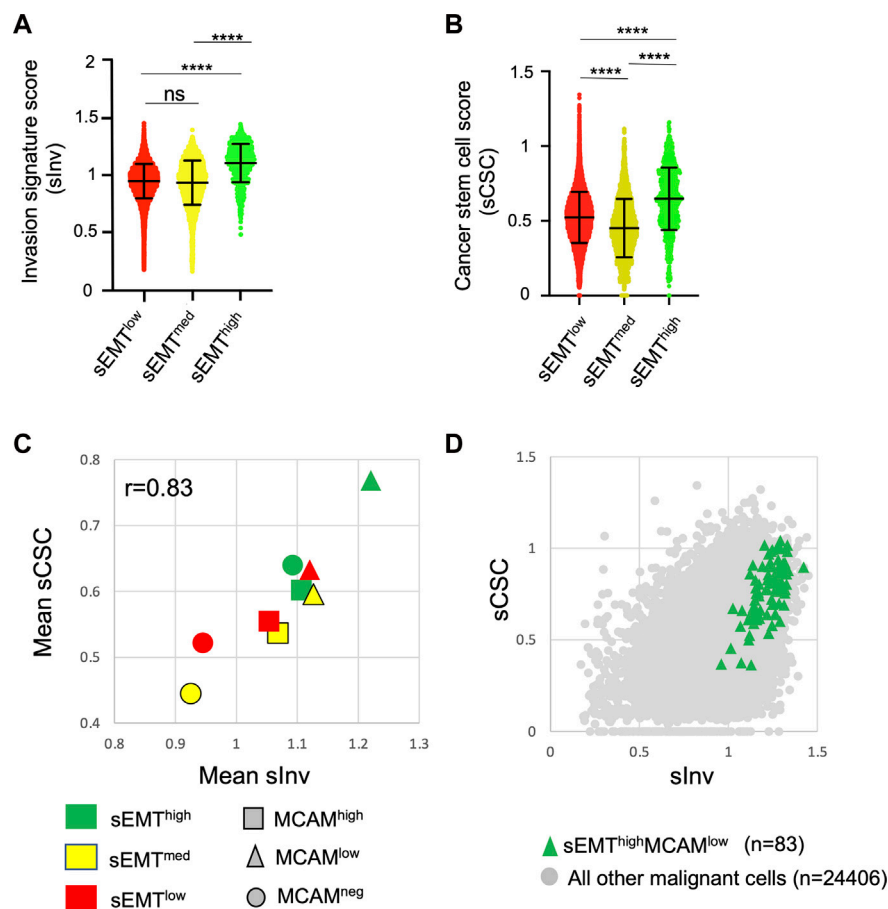
FIGURE 7
Heterogeneity of gene expression in malignant cells. **(A)** MCAM and EPCAM gene expression in single malignant cells (red; $n = 24,489$) and single normal epithelial cells (blue; $n = 4,355$), as identified by Wu et al. (2021). Percentages indicate the proportion of expressing cells (expression > 0). Differences in expression were analysed using a two-tailed Mann Whitney test; ****$p < 0.0001$. **(B)** MCAM gene expression as in **(A)**, except performed on four individual patients from the study. Anonymised patient IDs and the phenotype of their tumour (Wu et al., 2021) are shown along with the percentage of MCAM expressing cells in the malignant (red) and normal epithelium (blue). The number of malignant and normal epithelial cells in each tumour is also shown in red and blue respectively. Differences in expression were analysed using a two-tailed Mann Whitney test; **$p < 0.01$, ****$p < 0.0001$. **(C)** Population distribution of 22,489 malignant cells with respect to their EMT score (sEMT). The positions of the sEMT^{low}, sEMT^{med} and sEMT^{high} sub-populations are indicated, along with the number of cells in each sub-population. **(D)** Expression of EMT regulators in the sEMT^{low}, sEMT^{med} and sEMT^{high} sub-populations. Gene expression of the transcription factors OVOL2 and TWIST1 and the RNA splicing factors ESRP1 and QK1 are indicated. Data was analysed using a two-tailed Mann Whitney test; ****$p < 0.0001$. **(E)** Expression of the S100A4 gene in the sEMT^{low}, sEMT^{med} and sEMT^{high} populations. Data was analysed using a two tailed Mann Whitney test; ****$p < 0.0001$. **(F)** Expression of the MCAM gene in the sEMT^{low}, sEMT^{med} and sEMT^{high} populations. Data was analysed using a two-tailed Mann Whitney test; ****$p < 0.0001$. **(G)** Population distribution of 1,093 tumours (from TCGA) with respect to their EMT score (sEMT × 10^{-3}) The positions of the sEMT^{low}, sEMT^{med} and sEMT^{high} tumours are indicated, along with the numbers of tumours in each group. The sEMT groups were defined as follows; sEMT^{low}=sEMT<2,000, sEMT^{med}=sEMT<5,000, sEMT^{high}=sEMT>5,000. **(H)** Expression of the MCAM gene in the 1,093 breast cancer tumours from TCGA defined as sEMT^{low}, sEMT^{med} or sEMT^{high} tumours. Differences in expression were analysed using a two-tailed Mann Whitney test; **$p < 0.01$, ****$p < 0.0001$. **(I)** Correlation of MCAM gene expression and sEMT for the in the 1,093 breast cancer tumours from TCGA. Spearman's r, confidence intervals (and associated p-value) are indicated.

FIGURE 8

Invasive and cancer stem cell phenotypes of MCAM expressing cells. **(A)** Invasion signature score (sInv) of the sEMT[low], sEMT[med] and sEMT[high] populations. **(B)** Cancer stem cell score (sCSC) of the sEMT[low], sEMT[med] and sEMT[high] populations. For **(A,B)**, the graphs show the mean and standard deviation. Statistical analysis was performed using an unpaired, two-tailed Mann Whitney test; ****$p < 0.0001$, ns is not significant. **(C)** Mean sInv and sCSC of the sEMT[low], sEMT[med] and sEMT[high] populations further sub-divided according to MCAM gene expression. Cells with no detectable MCAM expression were identified and then MCAM expressing cells were divided into two equal sized groups of MCAM[low] or MCAM[high] cells. Spearman's r for the data is shown. **(D)** Invasion signature score (sInv) and cancer stem cell score (sCSC) of the sEMT[high]MCAM[low] sub-population ($n = 83$) compared to all other malignant cells ($n = 24,406$).

transcription factors which regulate EMT and for mRNA splicing factors which are differentially regulated in this differentiation pathway (Lambert and Weinberg, 2021). Importantly, these genes were not used to derive the sEMT. Expression of OVOL2, which represses EMT and thus favours the epithelial phenotype, was greatest in the sEMT[low] population and showed significantly decreasing expression in the sEMT[mid] and sEMT[high] cells. In contrast, expression of TWIST1, which favours the mesenchymal phenotype, increased significantly from sEMT[low] across the three populations. Similarly, expression of the epithelial splicing factor ESRP1 was significantly greater in the sEMT[low] cells, whereas the mesenchymal splicing factor QKI was greatest in the sEMT[high] cells (Figure 7D). The differential expression of these transcription and mRNA splicing factors validates the sEMT-based classification and suggests that the EMT[low] sub-population is enriched in epithelial-like cells and sEMT[high] is enriched for mesenchymal-like cells. The intermediate levels of expression of the transcription and splicing factors in sEMT[med] suggests that this population is

enriched in hybrid E/M cells (Pastushenko et al., 2018; Kröger et al., 2019; Celià-Terrassa and Jolly, 2020; Lüönd et al., 2021; Pastushenko et al., 2021). This is further supported by the substantial and significant gain in S100A4/FSP1 expression, a marker of mesenchymal cells (Ye et al., 2017), from sEMT[med] to sEMT[high] (Figure 7E). We analysed MCAM expression across these sub-populations and found that MCAM expressing cells were greatly enriched in the sEMT[high] population and expression levels increased significantly and progressively from sEMT[low] to sEMT[med] and sEMT[high] (Figure 7F). This result was confirmed using bulk tumour gene expression data (1,093 breast cancer patients from TCGA); there was a spectrum of sEMT across this cohort and, again, MCAM gene expression was highest in the sEMT[high] tumours and positively correlated with sEMT (Figures 7G–I). Furthermore, both MCAM gene expression and sEMT were positively correlated with TGFB1 gene expression in the TCGA cohort (Supplementary Figures S5A, B), consistent with the ability of TGF-β to induce EMT and MCAM gene expression (Lamouille et al., 2014; Ma et al., 2018). This data suggests that

heterogeneity of MCAM gene expression amongst the malignant epithelial cells in breast cancer results, at least in part, due to the EMT spectrum, both within and between tumours. However, patients from the TCGA cohort did not show statistically significant differences in overall survival when their tumours were classified according to sEMT (Supplementary Figure S5C). This failure to detect differences in survival is likely to result from the overlapping nature of the sEMT populations.

We attempted to address how levels of MCAM gene expression might be associated with the invasive and stem cell-like phenotypes that results from EMT. We derived an invasion score (sInv) and a cancer stem cell score (sCSC) for each of the ~25,000 malignant cells based on published breast cancer gene expression signatures (Patsialou et al., 2012; Pece et al., 2019). This showed that the sEMT^high population had significantly higher sInv and sCSC than the other populations (Figures 8A, B). We sub-divided the sEMT populations according to MCAM expression (high, low and no expression) and determined the mean sInv and sCSC for the nine sub-populations. Mean sInv and sCSC were strongly positively correlated, consistent with the co-acquisition of these phenotypes during EMT (Figure 8C; Spearman's r = 0.83). Not surprisingly, the nine sub-populations showed considerable overlap with respect to sInv and sCSC, but the small population of sEMT^high MCAM^low cells had the greatest mean combined sInv and sCSC, suggesting that breast cancer cells within this population have the greatest invasive and stem cell potential when MCAM is expressed at low levels (Figure 8D). This is consistent with our *in vitro* data, whereby reduction of MCAM expression in the MDA-MB-231 cells, a breast cancer cell line with mesenchymal characteristics (Taube et al., 2010; Zeng et al., 2012), enhanced its invasiveness. Separate comparisons showed that the sInv and sCSC of the sEMT^high MCAM^low population were significantly greater than other sub-populations (Supplementary Figures S6A, B). We investigated whether sEMT and MCAM expression levels affected patient outcomes using clinical data from the TCGA cohort. However, there were no significant differences in patient outcomes between MCAM^high and MCAM^low expressing tumours when analysing the three sEMT sub-populations (Supplementary Figure S7).

Discussion

Breast cancer metastasises to many sites, including bone, lung, liver and the brain and it is metastatic breast cancer that presents the major challenges to therapy (Weigelt et al., 2005). Our results demonstrate an inhibitory role for tumour cell-expressed CD146 in both the adhesion of breast cancer cells to EC and in migration through the endothelium. These tumour-EC interactions occur twice during metastasis, during extravasation of malignant cells from the primary tumour and again when circulating tumour cells intravasate and seed the metastasis.

Our analysis of patient transcriptome data reveals that MCAM gene expression is associated with the malignant phenotype, supporting previous findings from IHC studies of normal and malignant breast tissue (Shih et al., 1997; Chakraborty et al., 2006; Ouhtit et al., 2017). Furthermore, we show that a reduction in MCAM gene expression is associated with the increased methylation of the MCAM promoter in tumour tissue, an epigenetic modification that is known to repress gene

expression and highlighted as a key regulator of breast cancer progression (Pasculli et al., 2018). Previous breast cancer-focussed studies have demonstrated that ectopic expression of CD146 (in MCF-7 cells) suppresses tumour growth in a xenograft model (Shih et al., 1997) and that CD146 expression is inversely correlated with Matrigel invasion (Ouhtit et al., 2017). These data suggest that CD146 plays an inhibitory role in breast cancer progression, in keeping with the epigenetic silencing of the MCAM gene in tumour tissue. Paradoxically, increased MCAM gene expression in tumour tissue was associated with reduced survival. We suggest that these seemingly contradictory findings can be reconciled by considering the heterogeneity of MCAM gene expression in breast cancer and the plasticity of the tumour phenotype.

The CD146 molecule is highly expressed in endothelium where it plays an important role in regulating extravasation. Our analysis of MCAM gene expression at the single cell level in breast cancer confirmed high levels of MCAM expression in both EC and PVC. Markers of these cell types, together with the strong positive correlation between MCAM gene expression and an angiogenesis signature, suggest that total MCAM gene expression levels in bulk tumour material are dependent, in part, on the tumour vasculature. Angiogenesis is a defining feature of solid tumours and, in common with many other cancer types, increased vascularisation is associated with poor outcomes in breast cancer (Weidner et al., 1991; Horak et al., 1992; Hanahan and Weinberg, 2011).

As well as expression on vascular cells, CD146 is also expressed on the malignant cells of the tumour and de Kruijff et al. showed that high levels of CD146 on malignant breast cancer cells is indicative of poor patient outcomes (De Kruijff et al., 2018). Our data show that MCAM gene expression in breast cancer is indicative of the relative levels of EMT in the tumour sample. Thus, MCAM/CD146 expression is associated with poor prognosis in breast cancer in at least two ways, increased vascularisation of the tumour and as a marker of EMT of the malignant cells themselves.

The scRNA-seq dataset used in this study is derived from 26 breast tumours (Wu et al., 2021); we classified 922 of the 24,489 malignant cells as sEMT^high. However, all but 30 of these were from a single patient (ID4513). This TNBC sample was obtained post-chemotherapy, suggesting that sEMT^high cells might have been enriched due to their increased drug resistance, as has been previously demonstrated (Creighton et al., 2009). Although the sEMT^high population was largely represented by this single patient in the scRNA-seq dataset, 38 of the 1,093 breast tumours from the TCGA dataset were sEMT^high and this dataset showed a positive correlation between sEMT and MCAM gene expression. Expression of CD146 is induced by TGF-β and EMT in breast cancer cells and CD146 overexpression can drive EMT *in vitro* (Imbert et al., 2012; Zeng et al., 2012; Ma et al., 2018). When comparing EMT markers in CD146 expressing or non-expressing tumour cells, De Kruijff et al. (2018) found no link between CD146 expression and expression of EMT associated genes in breast tumours. However, our approach was different and we determined MCAM gene expression in cells and tumours based on their sEMT; this approach reveals that MCAM expression is highest in the more mesenchymal cells/tumours. Furthermore, our results show that heterogeneity of MCAM gene expression is found across the EMT spectrum. The epithelial phenotype predominates amongst the malignant cells and hence MCAM expressing epithelial cells are more numerous than

MCAM expressing mesenchymal cells. However, the levels of MCAM expression are significantly greater in the mesenchymal-like cells, with MCAM expression levels positively correlating with sEMT. It seems likely that greater numbers of malignant cells with higher levels of MCAM/CD146 expression is indicative of more cells undergoing EMT, this generates larger populations of cells with invasive and stem cell-like characteristics with a greater potential for metastasis and disease progression.

Our data reveal an inhibitory action of CD146 in breast cancer cell-EC adhesion and TEM, supporting previous work revealing that CD146 negatively regulates breast cancer progression (Shih et al., 1997; Chakraborty et al., 2006; Ouhtit et al., 2017). However, the CD146 molecule has previously been shown to play a positive role in the adhesion of melanoma cells to EC (Xie et al., 1997) and the TEM of monocytes and T cells (Bardin et al., 2009; Duan et al., 2013; Breuer et al., 2018). Furthermore, others have reported that CD146 plays a positive role in breast cancer cell migration, albeit using the parental MDA-MB-231 cell line (or other breast cancer cell lines) and not the brain metastatic variant of MDA-MB-231 (Zabouo et al., 2009; Imbert et al., 2012; Zeng et al., 2012). The basis for these differences in results are unclear. However, variation in CD146 expression levels (and the position of a particular isolate of a given cell line on the EMT spectrum) will likely influence the activity of CD146. This is exemplified by our results showing differential effects of CD146 knockdown in the parental and BrM variant, suggesting the importance of a threshold level of CD146 expression in determining inhibitory activity.

Gene signatures revealed that mesenchymal-like cells (EMT^{high}) had the highest invasive (sInv) and stemness (sCSC) scores, as expected given the well-established links between EMT, stemness and invasion. However, within the sEMT^{high} population, the MCAM^{low} expressing cells had greater sInv and sCSC than MCAM^{high} expressing cells. At first sight, these results seem at odds with literature showing that it is the hybrid E/M state that is associated with metastatic and cancer-initiating activity (Pastushenko et al., 2018; Kröger et al., 2019; Celià-Terrassa and Jolly, 2020; Deshmukh et al., 2021; Lüönd et al., 2021; Pastushenko et al., 2021). The sEMT^{med} population is enriched in hybrid E/M cells and this population might be expected to harbour the highest sInv and sCSC. However, the boundaries we used to define the different sEMT sub-populations are artificial and, as the population distributions show, there is not a clear distinction between these sEMT populations and there is substantial overlap with respect to sInv and sCSC. We suggest that sEMT^{high}MCAM^{low} population identified here includes cells with a hybrid E/M phenotype, whereas the sEMT^{high}MCAM^{high} cells contains fewer E/M hybrids and more fully mesenchymal cells. This overlap between the sub-populations as defined here by sEMT will weaken any association between the prevalence of these sub-populations in bulk tumours and patient outcomes, explaining the results of the survival analysis performed here. Despite this limitation it is possible to formulate a simple model of breast cancer progression in relation to sEMT and MCAM expression; malignant epithelial-like cells undergo EMT and during this process acquire the migratory, invasive and stem cell-like characteristics required for metastasis. The onset of EMT also induces CD146 expression; cells with lower CD146 expression and a metastatic phenotype express CD146 at a level insufficient to inhibit TEM and are more likely to metastasise. In contrast, TEM

will be inhibited by high CD146 expression in fully mesenchymal cells. Interestingly, CD146 is displayed on extracellular vesicles (EV) released by mouse breast cancer cells and this targets the EV to the lungs where they help to establish the premetastatic niche (Ghoroghi et al., 2021), with similar EV detected in patients with breast cancer (Ekström et al., 2022). This raises the interesting possibility that fully mesenchymal malignant cells may not metastasise themselves but provide a source of CD146 containing EV which contribute to the colonisation of metastatic sites by the hybrid E/M cells.

Importantly, CD146 is more than a marker of EMT and overexpression can drive EMT (Zeng et al., 2012). The MDA-MB-231 cell line used in our studies has a mesenchymal-like phenotype (Taube et al., 2010; Zeng et al., 2012) and it is possible that CD146 inhibition in this cell line pushes the phenotype towards the hybrid E/M state. Induction of MCAM expression during EMT appears to be inconsistent with the epigenetic silencing of MCAM gene expression in malignant cells compared to normal breast tissue. The precise relationships between these populations of cells remains to be determined. However, it seems plausible that the majority of malignant cells with an epithelial phenotype silence MCAM by promoter methylation leaving a minority of cells which induce MCAM expression upon EMT; this highlights the need to explore tumour heterogeneity in greater detail. Further complexity in the relationship between EMT and the hallmarks of cancer is illustrated by recent findings showing that EMT is not linear, but has branchpoints with alternative outcomes (Zhang et al., 2022).

The inhibitory activity of CD146 in adhesion and TEM were most pronounced in the brain metastasising variant MDA-BrM. We found that the enhanced CD146 expression on MDA-BrM was associated with a reduced TEM phenotype using HUVEC and hCMEC/D3 cells as a source of EC from the peripheral circulation and blood-brain barrier respectively. We speculated that MDA-BrM might demonstrate stronger binding to hCMEC/D3 than HUVEC and that the parental line would exhibit a preference for HUVEC, consistent with their tropism in *in vivo* models. Whilst there was some evidence of this selectivity at a single time point, this was not evident throughout the assay. Indeed, both the parental MDA and MDA-BrM derivative cell lines showed only weak binding to hCMEC/D3. This may reflect findings suggesting that adhesion to blood brain barrier EC is very weak in the absence of inflammation and that TEM at this site might be regulated differently to restrict the influx of immune cells into the brain (Engelhardt and Ransohoff, 2012). Alternatively, weak adhesion to hCMEC/D3 might reflect immortalisation by SV40/hTERT, resulting in differences between this cell line and primary blood brain barrier cells (Weksler et al., 2005; Urich et al., 2012; Weksler et al., 2013; Biemans et al., 2017).

In summary, expression of the CD146 molecule in breast cancer is of prognostic and functional importance. High levels of CD146 expression in bulk tumour reflect vascularisation and CD146 expression in the malignant cells is associated with EMT and increased invasive and stemness characteristics. Intermediate levels of MCAM gene expression are likely to be associated with the hybrid E/M state, whereas cells expressing high levels of CD146 are likely to be fully mesenchymal and have less metastatic activity *in vivo*. Our findings have relied extensively on informatics-based approaches using human breast cancer transcriptome profiles. Gene signatures underestimate the complexity of biological systems and, whilst they are valuable to infer phenotypes, they

are an imperfect approach. In addition, the use of cut-offs (e.g., in gene expression or signature scores) is arbitrary with respect to biological effects and it is important to now test these hypotheses and verify key findings in biological model systems. Our results demonstrate that understanding cellular and molecular heterogeneity in breast cancer is essential to understand and treat the underlying pathology.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

The study was conceived and designed by AM, AO, LM, PJ, and GC. AM and AO performed the *in vitro* experiments, LM, SB, and GC performed the *in silico* analyses and all authors contributed to the drafting and editing of the manuscript. Some of the data presented in this manuscript were reported in the PhD thesis of AM, submitted to the University of Leeds.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2023.1129015/full#supplementary-material

References

Bardin, N., Blot-Chabaud, M., Despoix, N., Kebir, A., Harhouri, K., Arsanto, J. P., et al. (2009). CD146 and its soluble form regulate monocyte transendothelial migration. *Arterioscler. Thromb. Vasc. Biol.* 29, 746–753. doi:10.1161/ATVBAHA.108.183251

Biemans, E. A. L. M., Jäkel, L., de Waal, R. M. W., Kuiperij, H. B., and Verbeek, M. M. (2017). Limitations of the hCMEC/D3 cell line as a model for Aβ clearance by the human blood-brain barrier. *J. Neurosci. Res.* 95, 1513–1522. doi:10.1002/jnr.23964

Breuer, J., Korpos, E., Hannocks, M. J., Schneider-Hohendorf, T., Song, J., Zondler, L., et al. (2018). Blockade of MCAM/CD146 impedes CNS infiltration of T cells over the choroid plexus. *J. Neuroinflammation* 15, 236. doi:10.1186/s12974-018-1276-4

Celià-Terrassa, T., and Jolly, M. K. (2020). Cancer stem cells and epithelial-to-mesenchymal transition in cancer metastasis. *Cold Spring Harb. Perspect. Med.* 10, a036905–a036917. doi:10.1101/cshperspect.a036905

Chakraborty, G., Rangaswami, H., Jain, S., and Kundu, G. C. (2006). Hypoxia regulates cross-talk between Syk and Lck leading to breast cancer progression and angiogenesis. *J. Biol. Chem.* 281, 11322–11331. doi:10.1074/jbc.M512546200

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18, 248–262. doi:10.1016/j.celrep.2016.12.019

Creighton, C. J., Li, X., Landis, M., Dixon, J. M., Neumeister, V. M., Sjolund, A., et al. (2009). Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc. Natl. Acad. Sci. U. S. A.* 106, 13820–13825. doi:10.1073/pnas.0905718106

De Kruijff, I. E., Timmermans, A. M., den Bakker, M. A., Trapman-Jansen, A. M. A. C., Foekens, R., Meijer-Van Gelder, M. E., et al. (2018). The prevalence of CD146 expression in breast cancer subtypes and its relation to outcome. *Cancers (Basel).* 10, 134. doi:10.3390/cancers10050134

Deshmukh, A. P., Vasaikar, S. V., Tomczak, K., Tripathi, S., den Hollander, P., Arslan, E., et al. (2021). Identification of EMT signaling cross-talk and gene regulatory networks by single-cell RNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2102050118. doi:10.1073/pnas.2102050118

Dillekås, H., Rogers, M. S., and Straume, O. (2019). Are 90% of deaths from cancer caused by metastases? *Cancer Med.* 8, 5574–5576. doi:10.1002/cam4.2474

Duan, H., Xing, S., Luo, Y., Feng, L., Gramaglia, I., Zhang, Y., et al. (2013). Targeting endothelial CD146 attenuates neuroinflammation by limiting lymphocyte extravasation to the CNS. *Sci. Rep.* 3, 1687. doi:10.1038/srep01687

Dudzik, P., Trojan, S. E., Ostrowska, B., Lasota, M., Dulińska-Litewka, J., Laidler, P., et al. (2019). Aberrant promoter methylation may be responsible for the control of CD146 (MCAM) gene expression during breast cancer progression. *Acta Biochim. Pol.* 66, 619–625. doi:10.18388/abp.2019_2907

Ekström, K., Crescitelli, R., Pétursson, H. I., Johansson, J., Lasser, C., and Olofsson Bagge, R. (2022). Characterization of surface markers on extracellular vesicles isolated from lymphatic exudate from patients with breast cancer. *BMC Cancer* 22, 50. doi:10.1186/s12885-021-08870-w

Engelhardt, B., and Ransohoff, R. M. (2012). Capture, crawl, cross: The T cell code to breach the blood-brain barriers. *Trends Immunol.* 33, 579–589. doi:10.1016/j.it.2012.07.004

Ghoroghi, S., Mary, B., Larnicol, A., Asokan, N., Klein, A., Osmani, N., et al. (2021). Ral GTPases promote breast cancer metastasis by controlling biogenesis and organ targeting of exosomes. *Elife* 10, 615399–e61629. doi:10.7554/eLife.61539

Györffy, B., Lanczky, A., Eklund, A. C., Denkert, C., Budczies, J., Li, Q., et al. (2010). An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* 123, 725–731. doi:10.1007/s10549-009-0674-9

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell* 144, 646–674. doi:10.1016/j.cell.2011.02.013

Horak, E. R., Leek, R., Klenk, N., LeJeune, S., Smith, K., Stuart, N., et al. (1992). Angiogenesis, assessed by platelet/endothelial cell adhesion molecule antibodies, as

indicator of node metastases and survival in breast cancer. *Lancet* 340, 1120–1124. doi:10.1016/0140-6736(92)93150-l

Imbert, A. M., Garulli, C., Choquet, E., Koubi, M., Aurrand-Lions, M., and Chabannon, C. (2012). CD146 expression in human breast cancer cell lines induces phenotypic and functional changes observed in epithelial to mesenchymal transition. *PLoS One* 7, e43752. doi:10.1371/journal.pone.0043752

Jing, X., Liang, H., Hao, C., Yang, X., and Cui, X. (2019). Overexpression of MUC1 predicts poor prognosis in patients with breast cancer. *Oncol. Rep.* 41, 801–810. doi:10.3892/or.2018.6887

Kröger, C., Afeyan, A., Mraz, J., Eaton, E. N., Reinhardt, F., Khodor, Y. L., et al. (2019). Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* 116, 7353–7362. doi:10.1073/pnas.1812876116

Lambert, A. W., Pattabiraman, D. R., and Weinberg, R. A. (2017). Emerging biological principles of metastasis. *Cell* 168, 670–691. doi:10.1016/j.cell.2016.11.037

Lambert, A. W., and Weinberg, R. A. (2021). Linking EMT programmes to normal and neoplastic epithelial stem cells. *Nat. Rev. Cancer* 21, 325–338. doi:10.1038/s41568-021-00332-6

Lamouille, S., Xu, J., and Derynck, R. (2014). Molecular mechanisms of epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* 15, 178–196. doi:10.1038/nrm3758

Li, Q., Yin, L., Jones, L. W., Chu, G. C. Y., Wu, J. B. Y., Huang, J. M., et al. (2016). Keratin 13 expression reprograms bone and brain metastases of human prostate cancer cells. *Oncotarget* 7, 84645–84657. doi:10.18632/oncotarget.13175

Li, Y., Ge, D., and Lu, C. (2019). The SMART App: An interactive web application for comprehensive DNA methylation analysis and visualization. *Epigenetics Chromatin* 12, 71. doi:10.1186/s13072-019-0316-3

Lüönd, F., Sugiyama, N., Bill, R., Bornes, L., Hager, C., Tang, F., et al. (2021). Distinct contributions of partial and full EMT to breast cancer malignancy. *Dev. Cell* 56, 3203–3221.e11. doi:10.1016/j.devcel.2021.11.006

Ma, Y., Zhang, H., Xiong, C., Liu, Z., Xu, Q., Feng, J., et al. (2018). CD146 mediates an E-cadherin-to-N-cadherin switch during TGF-β signaling-induced epithelial-mesenchymal transition. *Cancer Lett.* 430, 201–214. doi:10.1016/j.canlet.2018.05.016

Madsen, C. D., and Sahai, E. (2010). Cancer dissemination-lessons from leukocytes. *Dev. Cell* 19, 13–26. doi:10.1016/j.devcel.2010.06.013

Mamdouh, Z., Mikhailov, A., and Muller, W. A. (2009). Transcellular migration of leukocytes is mediated by the endothelial lateral border recycling compartment. *J. Exp. Med.* 206, 2795–2808. doi:10.1084/jem.20082745

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715. doi:10.1016/j.cell.2008.03.027

Mannion, A. J. (2022). Live cell imaging and analysis of cancer-cell transmigration through endothelial monolayers. *Methods Mol. Biol.* 2441, 329–338. doi:10.1007/978-1-0716-2059-5_26

Mannion, A. J., Odell, A. F., Taylor, A., Jones, P. F., and Cook, G. P. (2021). Tumour cell CD99 regulates transendothelial migration via CDC42 and actin remodelling. *J. Cell Sci.* 134, jcs240135. doi:10.1242/jcs.240135

McDermott, D. F., Huseni, M. A., Atkins, M. B., Motzer, R. J., Rini, B. I., Escudier, B., et al. (2018). Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. *Nat. Med.* 24, 749–757. doi:10.1038/s41591-018-0053-3

Moody, S. E., Perez, D., Pan, T. c., Sarkisian, C. J., Portocarrero, C. P., Sterner, C. J., et al. (2005). The transcriptional repressor Snail promotes mammary tumor recurrence. *Cancer Cell* 8, 197–209. doi:10.1016/j.ccr.2005.07.009

Onken, M. D., Li, J., and Cooper, J. A. (2014). Uveal melanoma cells utilize a novel route for transendothelial migration. *PLoS One* 9, e115472. doi:10.1371/journal.pone.0115472

Ouhtit, A., Abdraboh, M. E., Hollenbach, A. D., Zayed, H., and Raj, M. H. G. (2017). CD146, a novel target of CD44-signaling, suppresses breast tumor cell invasion. *Cell Commun. Signal.* 15, 45. doi:10.1186/s12964-017-0200-3

Pangeni, R. P., Channathodiyil, P., Huen, D. S., Eagles, L. W., Johal, B. K., Pasha, D., et al. (2015). The GALNT9, BNC1 and CCDC8 genes are frequently epigenetically dysregulated in breast tumours that metastasise to the brain. *Clin. Epigenetics* 7, 57. doi:10.1186/s13148-015-0089-x

Park, S. J., Yoon, B. H., Kim, S. K., and Kim, S. Y. (2019). GENT2: An updated gene expression database for normal and tumor tissues. *BMC Med. Genomics* 12 (5), 101. doi:10.1186/s12920-019-0514-7

Pasculli, B., Barbano, R., and Parrella, P. (2018). Epigenetics of breast cancer: Biology and clinical implication in the era of precision medicine. *Semin. Cancer Biol.* 51, 22–35. doi:10.1016/j.semcancer.2018.01.007

Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., et al. (2018). Identification of the tumour transition states occurring during EMT. *Nature* 556, 463–468. doi:10.1038/s41586-018-0040-3

Pastushenko, I., Mauri, F., Song, Y., de Cock, F., Meeusen, B., Swedlund, B., et al. (2021). Fat1 deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* 589, 448–455. doi:10.1038/s41586-020-03046-1

Patsialou, A., Wang, Y., Lin, J., Whitney, K., Goswami, S., Kenny, P. A., et al. (2012). Selective gene-expression profiling of migratory tumor cells *in vivo* predicts clinical outcome in breast cancer patients. *Breast Cancer Res.* 14, R139. doi:10.1186/bcr3344

Pece, S., Disalvatore, D., Tosoni, D., Vecchi, M., Confalonieri, S., Bertalot, G., et al. (2019). Identification and clinical validation of a multigene assay that interrogates the biology of cancer stem cells and predicts metastasis in breast cancer: A retrospective consecutive study. *EBioMedicine* 42, 352–362. doi:10.1016/j.ebiom.2019.02.036

Piali, L., Hammel, P., Uherek, C., Bachmann, F., Gisler, R. H., Dunon, D., et al. (1995). CD31/PECAM-1 is a ligand for αvβ3 integrin involved in adhesion of leukocytes to endothelium. *J. Cell Biol.* 130, 451–460. doi:10.1083/jcb.130.2.451

Rahn, J. J., Chow, J. W., Horne, G. J., Mah, B. K., Emerman, J. T., Hoffman, P., et al. (2005). MUC1 mediates transendothelial migration *in vitro* by ligating endothelial cell ICAM-1. *Clin. Exp. Metastasis* 22, 475–483. doi:10.1007/s10585-005-3098-x

Reymond, N., D'Água, B. B., and Ridley, A. J. (2013). Crossing the endothelial barrier during metastasis. *Nat. Rev. Cancer* 13, 858–870. doi:10.1038/nrc3628

Reymond, N., Im, J. H., Garg, R., Vega, F. M., Borda d'Agua, B., Riou, P., et al. (2012a). Cdc42 promotes transendothelial migration of cancer cells through β1 integrin. *J. Cell Biol.* 199, 653–668. doi:10.1083/jcb.201205169

Reymond, N., Imbert, A. M., Devilard, E., Fabre, S., Chabannon, C., Xerri, L., et al. (2004). DNAM-1 and PVR regulate monocyte migration through endothelial junctions. *J. Exp. Med.* 199, 1331–1341. doi:10.1084/jem.20032206

Reymond, N., Riou, P., and Ridley, A. J. (2012b). Rho GTPases and cancer cell transendothelial migration. *Methods Mol. Biol.* 827, 123–142. doi:10.1007/978-1-61779-442-1_9

Schenkel, A. R., Mamdouh, Z., Chen, X., Liebman, R. M., and Muller, W. A. (2002). CD99 plays a major role in the migration of monocytes through endothelial junctions. *Nat. Immunol.* 3, 143–150. doi:10.1038/ni749

Shih, I. M., Hsu, M. Y., Palazzo, J. P., and Herlyn, M. (1997). The cell-cell adhesion receptor Mel-CAM acts as a tumor suppressor in breast carcinoma. *Am. J. Pathol.* 151, 745–751.

Taube, J. H., Herschkowitz, J. I., Komurov, K., Zhou, A. Y., Gupta, S., Yang, J., et al. (2010). Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl. Acad. Sci. U. S. A.* 107, 15449–15454. doi:10.1073/pnas.1004900107

Urich, E., Lazic, S. E., Molnos, J., Wells, I., and Freskgård, P. O. (2012). Transcriptional profiling of human brain endothelial cells reveals key properties crucial for predictive *in vitro* blood-brain barrier models. *PLoS One* 7, e38149. doi:10.1371/journal.pone.0038149

van der Gun, B. T. F., Melchers, L. J., Ruiters, M. H. J., de Leij, L. F. M. H., McLaughlin, P. M. J., and Rots, M. G. (2010). EpCAM in carcinogenesis: The good, the bad or the ugly. *Carcinogenesis* 31, 1913–1921. doi:10.1093/carcin/bgq187

Varešlija, D., Priedigkeit, N., Fagan, A., Purcell, S., Cosgrove, N., O'Halloran, P. J., et al. (2019). Transcriptome characterization of matched primary breast and brain metastatic tumors to detect novel actionable targets. *J. Natl. Cancer Inst.* 111, 388–398. doi:10.1093/jnci/djy110

Varley, K. E., Gertz, J., Roberts, B. S., Davis, N. S., Bowling, K. M., Kirby, M. K., et al. (2014). Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res. Treat.* 146, 287–297. doi:10.1007/s10549-014-3019-2

Vestweber, D. (2015). How leukocytes cross the vascular endothelium. *Nat. Rev. Immunol.* 15, 692–704. doi:10.1038/nri3908

Wang, H. S., Hung, Y., Su, C., Peng, S., Guo, Y., Lai, M., et al. (2005). CD44 Cross-linking induces integrin-mediated adhesion and transendothelial migration in breast cancer cell line by up-regulation of LFA-1 (αLβ2) and VLA-4 (α4β1). *Exp. Cell Res.* 304, 116–126. doi:10.1016/j.yexcr.2004.10.015

Wang, Z., and Yan, X. (2013). CD146, a multi-functional molecule beyond adhesion. *Cancer Lett.* 330, 150–162. doi:10.1016/j.canlet.2012.11.049

Weidner, N., Semple, J. P., Welch, W. R., and Folkman, J. (1991). Tumor angiogenesis and metastasis — correlation in invasive breast carcinoma. *N. Engl. J. Med.* 324, 1–8. doi:10.1056/NEJM199101033240101

Weigelt, B., Peterse, J. L., and Van't Veer, L. J. (2005). Breast cancer metastasis: Markers and models. *Nat. Rev. Cancer* 5, 591–602. doi:10.1038/nrc1670

Weksler, B. B., Subileau, E. A., Perriere, N., Charneau, P., Holloway, K., Leveque, M., et al. (2005). Blood-brain barrier-specific properties of a human adult brain endothelial cell line. *FASEB J.* 19, 1872–1874. doi:10.1096/fj.04-3458fje

Weksler, B., Romero, I. A., and Couraud, P. O. (2013). The hCMEC/D3 cell line as a model of the human blood brain barrier. *Fluids Barriers CNS* 10, 16. doi:10.1186/2045-8118-10-16

Wu, S. Z., Al-Eryani, G., Roden, D. L., Junankar, S., Harvey, K., Andersson, A., et al. (2021). A single-cell and spatially resolved atlas of human breast cancers. *Nat. Genet.* 53, 1334–1347. doi:10.1038/s41588-021-00911-1

Xie, S., Luca, M., Huang, S., Gutman, M., Reich, R., Johnson, J. P., et al. (1997). Expression of MCAM/MUC18 by human melanoma cells leads to increased tumor growth and metastasis. *Cancer Res.* 57, 2295–2303.

Ye, X., Brabletz, T., Kang, Y., Longmore, G. D., Nieto, M. A., Stanger, B. Z., et al. (2017). Upholding a role for EMT in breast cancer metastasis. *Nature* 547, E1–E3. doi:10.1038/nature22816

Yoneda, T., Williams, P. J., Hiraga, T., Niewolna, M., and Nishimura, R. (2001). A bone-seeking clone exhibits different biological properties from the MDA-MB-231 parental human breast cancer cells and a brain-seeking clone *in vivo* and *in vitro*. *J. Bone Min. Res.* 16, 1486–1495. doi:10.1359/jbmr.2001.16.8.1486

Yu, L. G., Andrews, N., Zhao, Q., McKean, D., Williams, J. F., Connor, L. J., et al. (2007). Galectin-3 interaction with Thomsen-Friedenreich disaccharide on cancer-associated MUC1 causes increased cancer cell endothelial adhesion. *J. Biol. Chem.* 282, 773–781. doi:10.1074/jbc.M606862200

Zabouo, G., Imbert, A. M., Jacquemier, J., Finetti, P., Moreau, T., Esterni, B., et al. (2009). CD146 expression is associated with a poor prognosis in human breast tumors

and with enhanced motility in breast cancer cell lines. *Breast Cancer Res.* 11, R1. doi:10.1186/bcr2215

Zen, K., Liu, D. Q., Guo, Y. L., Wang, C., Shan, J., Fang, M., et al. (2008). CD44v4 is a major E-selectin ligand that mediates breast cancer cell transendothelial migration. *PLoS One* 3, e1826. doi:10.1371/journal.pone.0001826

Zeng, Q., Li, W., Lu, D., Wu, Z., Duan, H., Luo, Y., et al. (2012). CD146, an epithelial-mesenchymal transition inducer, is associated with triple-negative breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1127–1132. doi:10.1073/pnas.1111053108

Zhang, Y., Donaher, J. L., Das, S., Li, X., Reinhardt, F., Krall, J. A., et al. (2022). Genome-wide CRISPR screen identifies PRC2 and KMT2D-COMPASS as regulators of distinct EMT trajectories that contribute differentially to metastasis. *Nat. Cell Biol.* 24, 554–564. doi:10.1038/s41556-022-00877-0