Neurogenesis and angiogenesis are associated features of aggressive breast cancer

Elisabeth Wik[1,2#], Sura Aziz[1,2#], Heidrun Vethe[1], Cecilie Askeland[1,2], Kenneth Finne[1], Gøril Knutsvik[1,2], Ingeborg Winge[1], Kristi Krüger[1], Amalie A. Svanøe[1], Even Birkeland[1,3], Silje Kjølle[1], Benedicte Davidsen[4], Turid Aas[4], Ingunn M. Stefansson[1,2], Dimitrios Kleftogiannis[1,5], Lars A. Akslen[*1,2]

[#] *These authors contributed equally*

[1] Centre for Cancer Biomarkers CCBIO, Department of Clinical Medicine, University of Bergen, Bergen, Norway.
[2] Department of Pathology, Haukeland University Hospital, Bergen, Norway.
[3] Department of Biomedicine, University of Bergen, Bergen, Norway
[4] Department of Surgery, Haukeland University Hospital, Bergen, Bergen, Norway.
[5] Computational Biology Unit (CBU), Department of Informatics, University of Bergen, Bergen, Norway

---

*** Corresponding author:**

Prof. Lars A. Akslen, MD PHD
Centre for Cancer Biomarkers CCBIO
Department of Clinical Medicine
Section for Pathology, University of Bergen,
Bergen, Norway
Email: lars.akslen@uib.no
Tel: +4755973182

Conflict of interest statement:

The authors have declared that no conflict of interest exists.

Abstract

The tumor microenvironment (TME) is important for cancer growth and progress. Whereas angiogenesis is an established hallmark of cancer, the field of cancer neuroscience is now emerging. Here, we asked whether neurogenesis and angiogenesis are associated features in breast cancer and related to aggressive behavior. Single cell-based spatial analysis by imaging mass cytometry (IMC) indicated significant colocalization of neural and vascular structures, suggesting the presence of neuro-vascular niches within the tumor tissue. By protein and gene expression data from human breast cancer cohorts, we found that neurogenesis and angiogenesis, together with a consolidated neuro-angiogenic signature score, were linked to high-grade breast cancer characteristics and reduced survival, also within the luminal tumor subgroup. Cases with high neuro-angiogenic score were split in two subgroups, one hormone receptor negative and one receptor positive. Our findings might be relevant for improved patient stratification and further exploration of novel treatment strategies.

Introduction

Tumor-stroma communication is important for cancer progression and is suggested as a therapeutic target. Programs of inflammation, angiogenesis, and fibroblast activation are among the tumor microenvironment (TME) characteristics of aggressive cancer (1,2). Recently, evidence has indicated links between tumor cells, nerve elements and progressive disease (3,4), and experimental studies suggest that cancer cells may secrete certain factors that simultaneously promote the recruitment of both sprouting axons and endothelial cells (5-9). Studies of neural tissue in breast cancer stroma are limited (10-16), and the potential coexistence and coregulation of microaxons and microvessels have, to our knowledge, not been examined in human tumor tissues.

Here, we investigated whether neurogenesis (*i.e.,* axonogenesis) and angiogenesis are associated features in breast cancer TME, and the relationship between these structures and aggressive cancer phenotypes were examined. Using single cell-based spatial analysis by imaging mass cytometry (IMC), together with protein and gene expression data from human breast cancer cohorts, our tissue-based findings indicate that neurogenesis and angiogenesis are coexisting and appear to have an impact on disease progression. Our findings provide a basis for better tumor classification and improved patient management.

Results

Markers of neurogenesis and angiogenesis associate with aggressive breast cancer

Our purpose was to study the presence and significance of neurogenesis (*i.e.*, presence of nerve fibers) and angiogenesis in human breast cancer, their association, and how they are linked to clinico-pathologic phenotypes and patient prognosis. We

also wanted to examine the spatial relationship between these components in the TME. We first applied immunohistochemistry (IHC) on tissues from 534 primary breast cancers and 95 matched lymph node metastases (Bergen Breast Cancer Cohort).

Neurogenesis in primary tumors and lymph node metastases. Among 483 primary tumors with sufficient tissue (90%), microaxons were detected in 102 cases (21%; **Figure 1a**) and nerve bundles in 126 cases (36%; **Figure 1b**). Summary statistics revealed a median MAD (microaxon density) of 0.058 per mm^2 (range 0–1.06), and median NBD (nerve bundle density) of 0.080 per mm^2 (range 0–0.10). Microaxons (presence vs. absence) associated with large tumor diameter, high histologic grade, ER negativity, high tumor cell proliferation by Ki67 expression, and molecular subtype; higher MAD was observed among HER2 positive and triple negative tumors (**Table 1**). No associations were observed between NBD and primary tumor features.

In lymph node (LN) metastases (n=95), 23 micrometastases and 14 additional cases were excluded due to insufficient material. Microaxons were present in 36 of 58 metastatic lymph nodes (62%), with MAD exhibiting reduced values compared to matched primary tumors (P=0.001). Presence of microaxons in LN metastases was associated with larger primary tumor diameter and negative ER and PR. Again, the aggressive molecular subtypes showed the highest values (*i.e.,* HER2+ and triple negative (borderline, p=0.07) subgroups; **Supplementary Table 1**). Nerve bundles were not detected in lymph nodes with metastases.

Angiogenesis in primary tumors and lymph node metastases. We then examined markers of angiogenesis, *i.e.,* microvessel density (MVD, vessel count/mm^2), vascular proliferation (pMVD, count of proliferating vessels/mm^2) and VPI (vascular proliferation index; pMVD/MVD x 100). Among 461 primary tumors (86%) (**Figure 1c-d**), summary

statistics revealed median pMVD of 3.07 (range 0-37.8/mm^2), median VPI of 3.5 % (range 0-49.0), and median MVD of 77.6 (range 13.9-318.0/mm^2). High pMVD and VPI scores showed consistent associations with features of aggressive tumors; scores were higher in HER2 positive, triple negative and basal-like subtypes (**Table 2; Supplementary Figure 1a-b**).

Results for LN metastases paralleled the primary tumor tissues (**Supplementary Table 2**). Angiogenesis by pMVD and VPI scores were reduced in LN metastases compared to matched primary tumors (**Supplementary Table 3**); VPI decreased within all molecular subtypes (**Supplementary Table 4**).

Markers of neurogenesis and angiogenesis are associated traits in breast cancer tissues

By case-based correlation analysis, neurogenesis (by MAD count) and angiogenesis (by pMVD count and VPI score) were positively correlated in primary tumors (for both: r(s)=0.3, P<0.001); no correlation was detected between MAD and MVD counts (**Figure 1e**; **Supplementary Figure 2a-c**). Higher VPI scores were observed in primary tumors with presence of microaxons (P<0.001; **Figure 1f**); this correlation was observed in luminal/HER2 negative and HER2 positive subtypes (**Figure 1g-k**). In contrast, no associations were observed between nerve bundles and vascular markers. Taken together, our IHC studies revealed positive correlations between neurogenesis and angiogenesis in primary breast cancer tissues.

Spatial analysis of neural and vascular structures indicates co-localization in breast cancer tissues

To investigate the single-cell spatial distribution and potential co-localization of neural and vascular structures in breast cancer, we designed an imaging mass cytometry

(IMC) panel of 38 markers; 12 of these markers were included in the current analyses (**Supplementary Table 5**) to profile 10 samples from luminal-like (n=5) and basal-like (n=5) tumors. IMC uses isotope-labelled antibodies combined with laser ablation and mass-spectrometry to produce high dimensional images which are further segmented into single-cell information to investigate co-expression and spatial distribution of various markers. Altogether, we identified 46,574 single-cells or equally denoted cell units (*i.e.*, 25,532 from basal-like and 21,042 from luminal-like tissues), quantifying the marker expression together with spatial information of each cell unit. An unsupervised clustering approach was used to identify cell phenotypic clusters of different epithelial phenotypes, T-cells, B-cells, macrophages, endothelial cells, and stromal cells (17). Vascular proliferation was determined by dividing endothelial cells on the basis of Ki67 and CD31 co-expression, using density estimates from the flowDensity software (18), resulting in the following four subsets (denoted as vascular phenotypes): Ki67-low/CD31-high, Ki67-high/CD31-high, Ki67-high/CD31-low and Ki67-low/CD31-low (**Supplementary Figure 3a**). Quantification of cell units in each subset (**Figure 2a**) indicated that higher number of proliferating vessels was present in basal-like compared to luminal-like tumors (69.03% vs. 30.97%). Our analysis showed that putative pre-existing and not proliferating vessels were present in both subgroups, with luminal-like cancers having a higher relative percentage.

To elucidate whether markers of neural fibers and vasculature are co-expressed within the same cell units (or within overlapping and densely packed units), we used density estimation of the Neurofilament marker that resulted in groups of low and high expression (**Supplementary Figure 3b**). Our analyses revealed that cell units of higher vascularization (CD31-high) exhibited higher levels of neural fiber elements based on Neurofilament expression ($P<0.001$; **Figure 2b-c**). Thus, using median

expression values of CD31 and Neurofilament across the studied cell units, our analysis indicated consolidated neurogenic and angiogenic features, *i.e.* the existence of a “neuro-angiogenic profile” (**Figure 2d**). We then asked whether the consolidated neuro-angiogenic profile could be used to discriminate between basal-like and luminal-like cases at the sample level. To do so, we used as input the consolidated neuro-angiogenic profile and projected basal-like and luminal-like samples into 2D space, using Multidimensional Scaling (MDS). We found that the basal-like samples were projected in closer proximity to each other, and they were clearly separated from the luminal-like samples (**Figure 2e).**

The identified neuro-angiogenic profiles or units were also detected visually, when cell units were mapped to the 2D space using t-distributed stochastic neighbor embedding (TSNE) (**Figure 3a)**. Further, we used IMC-based images to detect areas with spatial co-localization of neural and vascular structures. For each endothelial cell unit, we estimated the number of neighboring cells using a k-nearest approach, and we visualized the average distance from all cell units with high Neurofilament marker expression. Our results indicated that proliferating vessels (Ki67-high/CD31-high) were on average in closer proximity to cell units of high neural participation (by Neurofilament expression; Kolmogorov–Smirnov statistic; **Figure 3b**), when compared with the rest. Our results were validated using pseudo-colored IMC images with representative markers indicating proximity between neural and vascular structures (**Figure 3c-d**).

Taken together, our single-cell spatial analysis revealed that markers reflecting neurogenesis (Neurofilament) and angiogenesis (CD31), are co-expressed and/or co-localized, supporting relationships which might be functionally important. Thus, our

single cell-based IMC data are in concordance with observations from our case-based IHC analyses.

Proteomic analysis indicates differential expression of neurogenic and angiogenic proteins in breast cancer tissues

We then used mass spectrometry profiling of 24 laser microdissected tumor epithelial samples from basal-like (n=12) and luminal-like (n=12) breast cancer tissues to understand whether neurogenic and angiogenic proteins are differentially expressed in luminal and basal cell types, Gene Ontology (GO) annotations were used to identify gene sets related to axonogenic (GO:0007409), neurogenic (GO:0022008) and angiogenic (GO:0001525) processes. Due to the high degree of overlap between axonogenic and neurogenic annotations from GO, these two sets were merged and are referred to as “neurogenic” in the following. After duplicate removal, we compiled a set of 473 genes characterizing axon morphogenesis, axon growth, neural process generation, nervous system cell generation, neural cell differentiation, and another set of 494 genes characterizing blood vessel formation from pre-existing vessels.

In our proteomic dataset of 4122 quantified proteins, we detected 102 proteins from the neurogenesis set, and 121 proteins from the angiogenesis set. Bioinformatics analysis revealed 967 differentially enriched proteins between basal-like and luminal-like cases (**Figure 4a**); 25 proteins from the neurogenesis set, and 17 proteins from the angiogenesis set (hypergeometric probability of 0.01). Focusing on differentially enriched proteins from the neurogenic set, S100A6 exhibited the highest log-fold change (~3.0) between basal-like and luminal-like types, whereas MAPT exhibited the lowest log-fold change (~-5.0) (**Supplementary Figure 4a; Supplementary Table 6a**)**.** Focusing on differentially enriched proteins from the angiogenic set, HK2 exhibited the highest log-fold change (~3.4) between basal-like and luminal-like types, whereas

AGO1 exhibited the lowest log-fold change (~-2.1) (**Supplementary Figure 4b; Supplementary Table 6b**). Collectively, our proteomics data indicates that several proteins related to neurogenesis and angiogenesis are differentially expressed between basal-like and luminal-like subtypes.

We hypothesized that the identified protein expression profiles might characterize aggressive tumor features. To investigate this, we projected basal-like and luminal-like samples into a two-dimensional space, using MDS. Basal-like samples were projected in closer proximity to each other. Similarly, the luminal-like samples were grouped together, occupying a different space from the basal-like samples, resulting in two sample clusters (R^2~0.98; **Figure 4**). Taken together, the results from proteomics profiling are in concordance with our findings from IHC and IMC, providing evidence of neurogenic and angiogenic features characterizing differences between basal-like and luminal-like subtypes and aggressive breast cancer subgroups.

Gene expression analysis support neurogenesis and angiogenesis as features of aggressive breast cancer

Following up on our results from IHC, IMC and proteomics analyses, we wanted to elucidate how gene expression patterns representing sprouting axons and vascular proliferation relate to breast cancer phenotypes, by exploring gene expression programs reflecting tumor progression. By analyses of global gene expression data from human breast cancer in METABRIC (n=1710) and TCGA (n=505) databases, and breast cancer cell line data from CCLE (n= 59) and in-house cell lines (n=12), we mapped gene expression signatures reflecting stroke-associated sprouting axons (19) and vascular proliferation (20), and we then estimated the sprouting axons score (SAS) and vascular proliferation score (VPS; See Methods).

High SAS and VPS were both associated with ER and PR negative tumors (P<0.001), and their distribution across breast cancer subtypes revealed consistently higher scores in basal-like and HER2 enriched categories (**Figure 5a-b; Supplementary Figure 5a-b**), suggesting more neurogenesis and angiogenesis taking place in aggressive tumor subgroups.

We hypothesized an association between neurogenesis and angiogenesis in human breast cancer tissues. At the gene expression level, we found significant and strong correlations between the sprouting axons (SAS) and vascular proliferation scores (VPS) **(Figure 5c; Supplementary Figure 5c)**. Notably, similar and strong positive correlations were demonstrated using breast cancer cell-line data (CCLE and in-house), where basal-like cell-lines exhibited higher SAS and VPS compared to luminal-like cells (**Figure 5d; Supplementary Figure 6a**). Consequently, our findings indicated that neuro-angiogenic properties characterize aggressive breast cancer subtypes, in breast cancer tissues and cell-lines. These observations are in concordance with our findings from IHC and IMC. Collectively, our results support a coordination of neurogenesis and angiogenesis in aggressive breast cancer.

Next, we derived a consolidated score reflecting both sprouting axons and vascular proliferation, termed a neuro-angiogenic score, NAS (see Methods). We found a distinct separation between molecular subtypes, as also found in breast cancer cell lines; the highest scores were seen in basal-like and HER2 enriched subtypes (**Figure 5e-f; Supplementary Figure 5d, Supplementary Figure 6b**)**.** There was a positive correlation between NAS and expression of basal-like markers cytokeratin 5 and cytokeratin 14 in breast cancer cell lines (CCLE data; ρ=0.41 and 0.43, both P=0.001), indicating neuro-angiogenic features in association with aggressive breast cancer

subgroups, as found in both breast cancer tissues and cell lines, and concordant with our findings from IHC and IMC.

High neuro-angiogenic score associates with features promoting breast cancer progress. We then investigated enriched gene sets pointing to biological processes in NAS-high tumors. Signatures reflecting loss of estrogen receptor signaling, epithelial-mesenchymal transition (EMT), stemness, and hypoxia responses were among top ranked gene sets enriched in NAS-high tumors (GSEA; FDR<5%). We found down-regulation of ER signaling-related genes like ESR1, FOXA1, GATA3 (METABRIC and TCGA; SAM; FDR<0.005%; fold change < -5.3), adding to an association between high NAS and ER/PR negativity.

High neuro-angiogenic score associates with signatures reflecting tumor plasticity and stemness. Following up on the GSEA analyses demonstrating enriched gene sets reflecting EMT in NAS-high tumors, we demonstrated strong correlations between NAS and scores for EMT (21) and TGF-β (22) (**Supplementary Figure 7a-b; Supplementary Figure 8a-b**). As EMT is linked to stemness properties (23,24), and gene sets reflecting stemness were top ranked enriched in GSEA analyses, we assessed how NAS related to independent stemness signatures. We found strong positive correlations between NAS and a cancer stem cell score (25), and one novel Nestin-based stemness score (26) (**Supplementary Figure 7c-d; Supplementary Figure 8c-d**). We also observed strong correlations between NAS and mammary stem cell enriched and luminal progenitor signature scores (27), and strong negative correlation with the mature luminal signature score ($\rho = -0.58$; $P<0.001$).

Similarly, in CCLE breast cancer cell lines, NAS correlated with EMT, TGFβ and two stemness scores (**Supplementary Figure 9 a-d**). In summary, our findings support

relations between the neuro-angiogenic profile and features of plasticity and stemness, in breast cancer tissues and cancer cell lines.

High neuro-angiogenic score associates with signatures reflecting hypoxia responses. To further pursue our GSEA-based results, we found strong positive correlations between NAS and independent hypoxia signature scores (28,29), both in global tissue data and in cell lines (**Supplementary Figure 7e-f; Supplementary Figure 8e-f; Supplementary Figure 9 e-f**). These correlations were also significant within individual molecular subtypes. NAS significantly associated with presence of tumor necrosis, reflecting tumor hypoxia (TCGA cohort).

Phenotypic heterogeneity among tumors with increased neurogenesis and angiogenesis scores. When clustering dichotomized signature scores reflecting EMT, stemness, and hypoxia, along with phenotypic clinico-pathologic markers in accordance with high or low NAS, phenotypic subgroups among tumors were evident (four clusters, C1-C4; **Figure 6a-b)**. One subgroup of NAS-high tumors (C4) was ER negative and basal-like with high scores for EMT, TGF-β activation, stemness and hypoxia; the other subset of NAS-high cases (C3) was ER positive and non-basal but still associated with high EMT and stemness scores. No survival difference was observed between the two NAS-high subclusters (C3; C4).

The survival pattern for these clusters demonstrated best prognosis in cases of C1 and poorer survival in C2-C4 (**Figure 6c**), with no differences between C2-C4 (P>0.15). To summarize, we demonstrate phenotypic heterogeneity in the subset of tumors with NAS-high tumors, where ER status and hypoxia split the high-score cases.

The neuro-angiogenic tumor profile associates with reduced patient survival

As we demonstrated correlations between markers of sprouting axons and angiogenesis in breast cancer tissues and cell lines (protein and mRNA levels), and associations of these with aggressive breast cancer phenotypes, we next assessed how measures of neurogenesis and angiogenesis relate to breast cancer survival.

In our IHC-based tissue studies using Neurofilament positive microaxon densities (MAD) and nerve bundles (NBD) in primary tumors and lymph node metastases, limited prognostic information by survival analyses was found (**Figure 7a-b**). Still, presence of microaxons in lymph node metastases was associated with significantly shorter survival (**Figure 7c**). Angiogenesis markers (VPI, pMVD) in primary tumors were associated with shorter survival (**Figure 7d)**, whereas MVD was not significant. Vascular proliferation in lymph node metastases indicated reduced survival (p=0.05; **Figure 7e**). Notably, low VPI in primary tumors indicated improved survival within the basal-like category (p=0.05; **Figure 7f**). When including VPI in multivariate survival analysis, adjusting for traditional prognostic factors (tumor diameter, histologic grade, lymph node status), VPI maintained independent prognostic value (**Supplementary Table 7**). A separate multivariate model for the basal-like subgroup showed a trend of prognostic value for VPI (P=0.07).

Using METABRIC data, both high SAS and high VPS (dichotomized) were associated with reduced survival in univariate analyses (**Figure 7g-h)**. The consolidated neuro-angiogenic mRNA score (NAS) was associated with shorter survival, both in univariate and multivariate analysis, adjusting for tumor diameter, histologic grade, and lymph node status (**Figure 7i-j**). When adding PAM50-based molecular subtype to the multivariate model, NAS independently predicted reduced survival (METABRIC, HR=1.4; 95% CI 1.1-1.7; P=0.001). In addition, high NAS maintained independent association with survival within luminal tumors (HR=1.39; 95% CI 1.05-1.84; P=0.022),

with a similar trend in non-luminal tumors (P=0.09), when adjusting for tumor diameter, histologic grade, and lymph node status.

Discussion

The influence of the microenvironment on tumor progress has been increasingly recognized (30-32). Since the present topic has not been previously studied in human cancer, we asked whether neurogenesis and angiogenesis are associated traits in breast cancer and related to more aggressive behavior. Tissue-based protein markers and gene expression data from patient cohorts indicated that both neurogenesis and angiogenesis were consistently linked to high-grade tumors and reduced patient survival. By using spatial mapping at the single-cell level, neural and vascular structures were significantly colocalized. Our novel findings indicate that neurogenesis and angiogenesis might be coordinated and closely related in aggressive breast cancer.

In this integrated study, tumor-associated neurogenesis was strongly associated with markers of high-grade tumors, such as negative hormone receptors, tumor cell plasticity, stemness, and hypoxia responses. The highest level of microaxon density was found in triple negative and HER2 positive cancers, while the lowest density was observed in the luminal A subgroup. The findings are supported by recent data (15). Nerve bundles, assumed to be pre-existing structures, were observed in many primary tumors but did not associate with any high-grade traits. Notably, nerve bundles were not observed in lymph node metastases, where microaxons might possibly develop *de novo* from tissue or circulating cells (33). Patient survival was reduced in cases with presence of microaxons in lymph node metastases.

We then asked whether neurogenesis and angiogenesis are associated features in human breast cancer tissues. This was found to be the case, and a combined neuro-angiogenic mRNA-based score (NAS) was strongly linked to high-grade breast cancer and independently related to reduced patient survival by multivariate analysis. Notably, the neuro-angiogenic signature was even stronger than the tumor cell-based classification. Taken together, our data support a coordinated activation of neurogenesis and angiogenesis in human breast cancer.

Spatial mapping of single cells and exploration of potential neuro-vascular niches indicated significant proximity of neural and vascular structures, and this was more pronounced in basal-like as compared to luminal-like tumors. This colocalization, by single cell-based analysis using sensitive imaging mass cytometry, support our findings using case-based protein data and transcription profiles. Our results lend support to a functional relationship between neural and vascular components, as reflected in human breast cancer samples. Experimental studies have indicated commonalities for neuronal guidance and vascular sprouting, supporting that the two processes might be co-regulated (7,8), but this has not been previously addressed in human cancer tissues.

The consolidated neuro-angiogenic score was associated with markers of epithelial-mesenchymal plasticity, tumor stemness, and tumor hypoxia. These findings indicate that the connectivity of aggressive breast cancer features, based on tumor cell markers with characteristics of more invasive and less differentiated cells, are supplemented by varying degrees of “aggressive stroma”. Our data suggest an existence of neuro-vascular interactions in the TME which could potentially be targeted for more efficient cancer management.

Notably, when using multiple breast cancer cell lines (34), higher levels of the neuro-angiogenic score were observed in the most aggressive cancer subtypes. This indicates that neuro-angiogenic programs appear to be part of intrinsic tumor cell properties. The finding that our neuro-angiogenic score was prognostic within the luminal-like breast cancers, and the fact that some of the cases with high neuro-angiogenic score were hormone receptor positive, might indicate discordant evolution of tumor cells and TME in subsets of breast cancer.

Tumor angiogenesis is an established hallmark of cancer (1) and is well described in breast tumors (35,36). Here, we found consistent and strong associations between activated angiogenesis and aggressive breast cancer. Vascular proliferation was increased in basal-like cancers, as we have reported (36,37). Our previous findings were confirmed by the present single cell analysis using imaging mass cytometry. As a novel finding, low vascular proliferation within the subgroup of aggressive basal-like cancers indicated a better prognosis compared to the others. Also, elevated vascular proliferation in lymph node metastases was associated with reduced survival, which has not been previously reported.

In summary, our data support that neurogenesis and angiogenesis are associated features of aggressive breast cancer. High neuro-angiogenic score was consistently linked to high-grade breast cancer phenotypes and reduced patient survival. Notably, among some hormone receptor positive cases, a high-grade subgroup could be defined by increased neuro-angiogenic score. Single cell tissue mapping indicated colocalization of neural and vascular structures, suggesting the presence of neuro-vascular niches of possible functional importance. Taken together, our findings might be relevant for improved patient stratification and exploration of novel therapy.

Methods

Patient series

The study cohort included women diagnosed with primary invasive breast cancer as part of the prospective and population-based Norwegian Breast Cancer Screening Program during 1996-2003 (Hordaland County, Norway; 10% of the Norwegian population), age 50–69 years at time of diagnosis (median 59 years). Patients with distant metastatic disease at time of diagnosis (stage IV) were not included. From the whole cohort, nine patients refused to participate in the study, leaving 546 cases for inclusion; 403 screen-detected and 143 interval detected cancers. After exclusion of 12 cases due to lack of available tissue, 534 cases were finally included as previously described (38). The patients received treatment according to the national guidelines at the time (38). Detailed information about this study population and the specimen characteristics is presented in Supplementary Material.

Immunohistochemistry (IHC)

Sections from primary tumors and the matched axillary node with the largest metastatic focus (≥ 2 mm) were used for assessment of nerves (Neurofilament antibody). For Neurofilament staining (pan-nerve marker), the monoclonal mouse antibody Neurofilament (DAKO M0762) was used.

Sections from the same block of the primary tumor and the axillary lymph node with metastases that were used for assessment of tumor-related nerves, were applied for evaluation of angiogenesis markers; the polyclonal rabbit antibody (DAKO A0082) for Factor-VIII, and a monoclonal mouse antibody for Ki67 (DAKO 7240), were used. The blood vessels in the adjacent normal mammary or fat tissue stained with Factor-VIII and the proliferating tumor cells stained with Ki67 were used as internal controls for

these markers. In lymph nodes, the adjacent capillaries and the proliferating lymphoid cells were used as internal control cells for Factor-VIII and Ki67, respectively. The comparison with HE slides was always done to guide the location of primary tumors and metastatic foci in the breast and affected lymph nodes. For further details from the immunohistochemistry, see Supplementary Material.

Evaluation of tissue-based IHC for neurogenesis and angiogenesis markers

Microaxon and nerve bundle density. Microaxon and nerve bundle densities for each case were assessed separately in the whole tumor area using light microscopy (MAD=microaxon density; NBD=nerve bundle density). Microaxons were observed as single, small, and thin structures (**Figure 1a**), while a nerve bundle was identified as a group of axons (**Figure 1b**). The number of microaxons or bundles per mm^2 were determined.

Microvessel density (MVD). As a measure of angiogenesis, the count of all vessels in ten high power visual fields (HPF x400) was performed as described (35), and reported as counts/mm2. Microvessels included both vessel-like structures with a visible lumen and single endothelial cells or cell clusters **(Figure 1c)**, as defined by Factor-VIII positivity according to Weidner (39).

Proliferating microvessel density (pMVD). Using the same ten microscopic fields (x400) that were used for microvessel density counts, vessels containing Ki67 positive endothelial cells were counted. Thus, pMVD represents a separate count of vessels with proliferating endothelial cells and was reported as count/mm^2 (**Figure 1d**).

Vascular proliferation index (VPI). For estimation of activated angiogenesis, the proportion (%) of vessels with proliferating endothelial cells of the total number of microvessels counted in ten fields (Vascular proliferation index; VPI) was calculated.

Imaging mass cytometry (IMC)

Antibodies

The panel used in this study consisted of 35 metal-conjugated antibodies and two free metals, iridium, which binds double stranded DNA, and ruthenium, which binds tissue structures, seemingly non-specifically (40). Antibodies against cytokeratin (clones AE1/AE3) and CD31 (clone EPR3094) were purchased pre-conjugated (Fluidigm), while CD34 (clone ICO115; Cell signaling technologies) and neurofilament (clone C28E10; Abcam) were conjugated in-house using the Maxar X8 Multimetal Labeling Kit (Fluidigm). CD31 and CD34 were conjugated to the same metal; 151 Eu.

IMC staining protocol

Antibody hybridization was performed as described in the “Imaging Mass Cytometry Staining Protocol for FFPE Sectioned Tissue” (Fluidigm) with a few modifications. In short, the freshly cut TMA slide underwent dewaxing, rehydration and antigen retrieval in a Ventana Ultra Discovery Autostainer (Roche Diagnostics GmbH). The antibody mix was applied to the slide, which was then stored overnight at 4°C in a hydration chamber. After antibody incubation, the slide was first washed with PBS with and without detergent, before incubation first in 0.3 uM Iridium (Ir)-intercalator (Fluidigm), washed, and then in 0.0005% Ruthenium (RuO4)/PBS (Electron Microscopy Sciences) (40). Finally, the slides were washed in MaxPar H_2O, air-dried, and stored at 4 °C.

Antibody validation

All antibodies were validated by IHC using a test-TMA with positive control tissues for all antibodies included in the panel: tonsil, placenta, hippocampus, cerebellum, autonomic ganglion and peripheral nerve tissue, normal breast tissue, and selected breast carcinomas (ER+/PR+, HER2+, ER-/PR-/HER2- and basal-like breast carcinomas). In some cases, our IHC staining was compared to staining performed by

the Human Protein Atlas (proteinatlas.org). A pilot-TMA was also established and stained (five basal-like; five luminal-like).

IMC analysis and pre-processing

Data from the 10 pilot-TMA breast cancer cases were acquired by a Helios time-of-flight mass cytometer (CyTOF) coupled to a Hyperion Imaging System (Fluidigm) and administered using the CyTOF Software (v7.0.8493; Fluidigm). The square inscribed in the circular TMA cores were laser ablated at 200 Hz at a resolution of approximately 1μm^2.

The highly multiplexed data was processed using the ImcSegmentationPipeline(41) to segment data into spatial regions corresponding to single cells. First, the IMC output (.txt files) were converted to tiff-images (ImcTools v2.1.7), which were used for pixel classification (Ilastik v1.3.3post2) to generate probability maps of nuclei, cytoplasm/membrane, and background. Each probability map, corresponding to one TMA core, were then segmented into single cells, with output as a cell mask image, using Cell Profiler (v4.0.7). The pre-processed data were imported into Histocat (v1.7.6; (42)) and immediately exported as csv-files to be used in down-stream analyses. We note that even with high-quality segmentation, the imaging of tissue segments cannot rule out cases of “overlapping cell units” that do not capture the nucleus of an individual cell. Therefore, in a few cases nuclei-mismatched signal can be assigned to neighboring cells, especially in (cellular) dense areas.

IMC data normalization and analysis workflow

For each marker, the total intensity per cell unit was computed. Values were divided by the cell size and were arcsinh-transformed. The single-cell data per channel were censored at the 99th percentile to remove outliers. Single-cell annotation was

performed in a hierarchical scheme using unsupervised clustering/meta-clustering and prior-knowledge of cell type defining markers of our antibody panel, similar to the methodology presented in (17). The procedure operates in two steps: initially, single cells are categorized as Immune and Non-Immune. FlowSOM (43) is used to cluster the data into 120 clusters, that further merged based on cluster profile similarity. After this step, Phenograph (44) is used to refine the annotation and split Immune cells to three groups (B cells, T cells, macrophages), and non-Immune cells to three groups (epithelial cells, endothelial cells, stromal cells). For the clustering approach, data were standardized across markers using standard normalization.

To characterize vascular proliferation, we focused on the endothelial phenotype. All endothelial cells were gated into four subgroups (Ki67-low/CD31-high, Ki67-high/CD31-high, Ki67-high/CD31-low and Ki67-low/CD31-low) on the basis of CD31 and Ki-67 marker expression, using the flowDensity software (18). flowDensity is a supervised clustering algorithm based on density estimations that is designed to overcome the tedious work of setting static gates. flowDensity was also used to stratify single cells to Neurofilament high and Neurofilament low expression groups, given the density estimation of the Neurofilament marker, see Supplementary Material for further details on the evaluation of the tissue-based neurogenesis and angiogenesis markers.

For visualizations of cell units, high-dimensional single-cell data were mapped to 2D using the t-distributed stochastic neighbor embedding algorithm TSNE (45). We applied the Barnes–Hut implementation in R, with default parameters (seven initial dimensions: expression levels of Ki67, CD31, Neurofilament, Vimentin, aSMA, CD45, panCK; perplexity=30; θ=0.5).

For visualization of IMC samples to the 2D space, we used the Multidimensional Scaling (MDS) method (46) implemented in R with seven initial dimensions per sample: average marker expression of Ki67, CD31, Neurofilament, cell unit abundance (%) of Ki67-low/CD31-high, Ki67-high/CD31-high, Ki67-high/CD31-low and Ki67-low/CD31-low cells.

To estimate the distance of endothelial cells, to Neurofilament-high cells, we used a k-nearest neighbor approach. For every cell of the Ki67-low/CD31-high, Ki67-high/CD31-high, Ki67-high/CD31-low and Ki67-low/CD31-low phenotype, we first selected the k-closest cells. From the set of k-closest cells we filtered the ones with Neurofilament high expression, and we estimated the Euclidean distance from the corresponding endothelial cells of interest. The average log2-transformed distance was used for visualizations, and the procedure was repeated for different k values starting from 3 to 10 to assess reproducibility.

Proteomic analyses of microdissected breast cancer epithelial cells

Proteins were extracted using the filter-aided sample preparation (FFPE-FASP) protocol (47) and analyzed by liquid-chromatography tandem mass spectrometry (LC-MS/MS). Raw data were analyzed using the freely available Maxquant software (v1.6.0.16) with recommended settings for label-free quantification (48) (see Supplemental Material for an extended description of methods). Gene ontology datasets were collected from the AmiGO 2 resource (version 2.5.12). Protein intensities were compared between groups (basal-like and luminal-like) using the Students t-test. Correction for multiple hypothesis testing was done with the Benjamini-Hochberg method (49). Differential analysis of protein expression was performed using

the DEP R package (50). Volcano plots and other visualizations of the proteomics results were performed in R.

Gene expression (mRNA) analyses of breast cancer tissues and cell lines

In-house cell lines. Six basal-like and six luminal-like breast cancer cell lines were obtained from the American Type Culture Collection (ATCC, Manassas, VA). Total RNA from the cells was isolated with miRNeasy mini kit according to protocol (Qiagen, Venlo, NL). The global mRNA expression was examined by the Illumina Bead Array Technology (HumanHT-12 v4 Expression Bead Chip; Illumina, CA, USA). Microarray data were feature extracted using Genome Studio Software (Illumina, CA; USA), with default parameters with respect to the control categories (51) as part of the Whole-Genome Gene Expression Direct Hybridization Assay system. For further details about in-house cell lines and mRNA processing and analyses, see Supplementary Material.

External data sets. For the exploration of gene expression patterns related to neurogenesis and angiogenesis in breast cancer, publicly available microarray mRNA gene expression profiles, with information on clinico-pathologic and follow-up data, and molecular subtypes were analyzed. 1) the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohorts (discovery cohort, n=939 and validation cohort, n=845) (52); and The Cancer Genome Atlas (TCGA) breast cancer cohort (n=520) (53). Intrinsic molecular subtypes based on PAM50 classification (54) were available for all cohorts. Normal breast-like cases were excluded from all mRNA data sets. Further, we applied the breast cancer subset of the Cancer Cell Line Encyclopedia (CCLE) data, including 59 breast cancer cell lines (47 of these with the molecular subtype described) (34).

Sprouting axons and angiogenesis signatures. Previously published mRNA expression signature reflecting the transcriptional pattern of stroke-related sprouting axons in an experimental model of young mice (19), and a 32-gene signature (vascular proliferation score; VPS) previously associated with increased tissue-based microvessel proliferation (pMVD) in endometrial carcinomas (20) were mapped to the METABRIC and TCGA cohorts, and the CCLE data. For details about the signatures and the derivation of the corresponding signature scores, see Supplementary Material.

Consolidated neuro-angiogenesis profile. A combined neuro-angiogenic score (NAS), was derived by summarizing the sprouting axons and vascular proliferation scores.

Other gene expression signatures. For correlation analyses, we mapped two hypoxia signatures (28,29), an EMT signature (21), a TGF-β signature (22), and two signatures reflecting stemness features (25,26), to the mRNA data we analyzed. For further detailed information about gene expression analyses, see Supplementary Material.

Statistics and survival analyses

Data were analyzed using SPSS (Statistical Package of Social Scienc1es), Version 25.0 (Armonk, NY, USA; IBMM, Corp) and R-package via Rstudio interface version1.3.9. A two-sided P-value less than 0.05 was considered statistically significant. A P-value of 0.05-0.10 was considered to represent a statistical trend. Categories were compared using Pearson`s chi-square or Fisher`s exact tests when appropriate. Non-parametric correlations were tested by Spearman's rank correlation, while Mann-Whitney U and Kruskal-Wallis tests were used for comparing continuous variables in groups. Wilcoxon Signed Rank test was used to compare the differences between two continuous variables. The Kolmogorov–Smirnov statistic was used to quantify a distance between the between the empirical distribution functions of two

samples. Odds ratios (OR) and their 95% confidence intervals were calculated by the Mantel-Haenszel method. Kappa statistics were used to test inter- and intra-observer agreement of categorical data. For survival analyses, the endpoint was breast cancer specific survival, defined as the time in months from the date of histologic diagnosis to the date of death of breast cancer. Univariate survival analysis (Kaplan-Meier method) was performed using the log-rank test to compare differences in survival time between categories. Patients who died of other causes or who were alive at last date of follow-up were censored in the analyses. The influence of co-variates on breast cancer specific survival was analyzed by Cox`s proportional hazards method and tested by the backward stepwise likelihood ratio (lratio) test. All variables were tested by log-minus-log plots to determine their ability to be incorporated in multivariate modelling. When categorizing continuous variables, cut-off points were based on median or quartile values, also considering the distribution profile, the size of subgroups, and number of events in survival analyses.

Study approval

This study was approved by the Western Regional Committee for Medical and Health Research Ethics, REC West (REK 2014/1984). Written informed consent was not obtained from the patients, but in accordance with national ethics guidelines and procedures for retrospective studies, all participants were contacted with written information on the study and asked to respond if they objected. In total, 9 patients (1.7%) did not approve participation.

Author contribution

LAA conceived and designed the study and contributed to histologic analyses (IHC, IMC), data interpretation and in writing the major parts of the manuscript. EW contributed to the study design, analyzing the transcriptomic data, general data

analyses and interpretation, and in writing the major parts of the manuscript. SA contributed to the study design, performed the tissue-based work and IHC, performed statistical analysis, data interpretation and contributed to writing the manuscript. DK contributed to study design and analyzing the mRNA and IMC data, and contributed to writing the manuscript. HV and KF contributed to analyses of proteomics and IMC data. CA contributed to histologic analyses and work on IMC data, and to collection of clinico-pathologic data. IW contributed to IMC data. GK participated in data collection and interpretation. KK, AAS, EB, SK, and IMS participated in data collection and interpretation. TA and BD participated in collection of clinical data. All authors commented on and approved the final manuscript. EW and SA made the initial core findings, whereas EW played a larger role in manuscript completion; thus, EW is listed first in the authorship. All authors have read and approved the submitted paper.

Acknowledgements and funding information

We thank Gerd Lillian Hallseth and Bendik Nordanger for their excellent technical assistance. This work was partly supported by the Research Council of Norway through its Centre of Excellence funding scheme, project number 223250. The study was also supported by the Norwegian Cancer Society and Regional Health Trust Western Norway (Helse Vest).

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-674.
2. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 2022;12(1):31-46.
3. Hayakawa Y, et al. Nerve Growth Factor Promotes Gastric Tumorigenesis through Aberrant Cholinergic Signaling. *Cancer Cell.* 2017;31(1):21-34.
4. Monje M, et al. Roadmap for the Emerging Field of Cancer Neuroscience. *Cell.* 2020;181(2):219-222.
5. Mancino M, Ametller E, Gascon P, Almendro V. The neuronal influence on tumor progression. *Biochimica et biophysica acta.* 2011;1816(2):105-118.

6. Entschladen F, Palm D, Lang K, Drell TLt, Zaenker KS. Neoneurogenesis: tumors may initiate their own innervation by the release of neurotrophic factors in analogy to lymphangiogenesis and neoangiogenesis. *Med Hypotheses.* 2006;67(1):33-35.
7. Magnon C, et al. Autonomic nerve development contributes to prostate cancer progression. *Science.* 2013;341(6142):1236361.
8. Zahalka AH, et al. Adrenergic nerves activate an angio-metabolic switch in prostate cancer. *Science.* 2017;358(6361):321-326.
9. March B, et al. Tumour innervation and neurosignalling in prostate cancer. *Nature reviews Urology.* 2020;17(2):119-130.
10. Zhao Q, et al. The clinicopathological significance of neurogenesis in breast cancer. *BMC cancer.* 2014;14:484.
11. Lagadec C, et al. TrkA overexpression enhances growth and metastasis of breast cancer cells. *Oncogene.* 2009;28(18):1960-1970.
12. Reis-Filho JS, et al. Distribution and significance of nerve growth factor receptor (NGFR/p75NTR) in normal, benign and malignant breast tissue. *Mod Pathol.* 2006;19(2):307-319.
13. Tomellini E, et al. Nerve growth factor and proNGF simultaneously promote symmetric self-renewal, quiescence, and epithelial to mesenchymal transition to enlarge the breast cancer stem cell compartment. *Stem Cells.* 2015;33(2):342-353.
14. Huang D, et al. Nerve fibers in breast cancer tissues indicate aggressive tumor progression. *Medicine (Baltimore).* 2014;93(27):e172.
15. Jezequel P, et al. Identification of three subtypes of triple-negative breast cancer with potential therapeutic implications. *Breast Cancer Res.* 2019;21(1):65.
16. Pundavela J, et al. Nerve fibers infiltrate the tumor microenvironment and are associated with nerve growth factor production and lymph node invasion in breast cancer. *Mol Oncol.* 2015;9(8):1626-1635.
17. Keren L, et al. A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell.* 2018;174(6):1373-1387 e1319.
18. Malek M, Taghiyar MJ, Chong L, Finak G, Gottardo R, Brinkman RR. flowDensity: reproducing manual gating of flow cytometry data by automated density-based cell population identification. *Bioinformatics.* 2015;31(4):606-607.
19. Li S, et al. An age-related sprouting transcriptome provides molecular control of axonal sprouting after stroke. *Nat Neurosci.* 2010;13(12):1496-1504.
20. Stefansson IM, et al. Increased angiogenesis is associated with a 32-gene expression signature and 6p21 amplification in aggressive endometrial cancer. *Oncotarget.* 2015;6(12):10634-10645.
21. Jechlinger M, et al. Expression profiling of epithelial plasticity in tumor progression. *Oncogene.* 2003;22(46):7155-7169.
22. Padua D, et al. TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell.* 2008;133(1):66-77.
23. Liu X, Fan D. The epithelial-mesenchymal transition and cancer stem cells: functional and mechanistic links. *Current pharmaceutical design.* 2015;21(10):1279-1291.
24. Scheel C, Weinberg RA. Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links. *Semin Cancer Biol.* 2012;22(5-6):396-403.
25. Pece S, et al. Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell.* 2010;140(1):62-73.
26. Kruger K, et al. Expression of Nestin associates with BRCA1 mutations, a basal-like phenotype and aggressive breast cancer. *Sci Rep.* 2017;7(1):1089.
27. Lim E, et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med.* 2009;15(8):907-913.
28. Eustace A, et al. A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in laryngeal cancer but not bladder cancer. *Clin Cancer Res.* 2013;19(17):4879-4888.

29. Hu Z, et al. A compact VEGF signature associated with distant metastases and poor outcomes. *BMC Med.* 2009;7:9.
30. Maman S, Witz IP. A history of exploring cancer in context. *Nat Rev Cancer.* 2018;18(6):359-376.
31. Servick K. War of nerves. *Science.* 2019;365(6458):1071-1073.
32. Lambert AW, Pattabiraman DR, Weinberg RA. Emerging Biological Principles of Metastasis. *Cell.* 2017;168(4):670-691.
33. Mauffrey P, et al. Progenitors from the central nervous system drive neurogenesis in cancer. *Nature.* 2019;569(7758):672-678.
34. Barretina J, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature.* 2012;483(7391):603-607.
35. Arnes JB, et al. Vascular proliferation is a prognostic factor in breast cancer. *Breast Cancer Res Treat.* 2012;133(2):501-510.
36. Kruger K, Stefansson IM, Collett K, Arnes JB, Aas T, Akslen LA. Microvessel proliferation by co-expression of endothelial nestin and Ki-67 is associated with a basal-like phenotype and aggressive features in breast cancer. *Breast.* 2013;22(3):282-288.
37. Nalwoga H, Arnes JB, Stefansson IM, Wabinga H, Foulkes WD, Akslen LA. Vascular proliferation is increased in basal-like breast cancer. *Breast cancer research and treatment.* 2011;130(3):1063-1071.
38. Knutsvik G, et al. Evaluation of Ki67 expression across distinct categories of breast cancer specimens: a population-based study of matched surgical specimens, core needle biopsies and tissue microarrays. *PLoS One.* 2014;9(11):e112121.
39. Weidner N, Semple JP, Welch WR, Folkman J. Tumor angiogenesis and metastasis--correlation in invasive breast carcinoma. *N Engl J Med.* 1991;324(1):1-8.
40. Catena R, Montuenga LM, Bodenmiller B. Ruthenium counterstaining for imaging mass cytometry. *J Pathol.* 2018;244(4):479-484.
41. Zanotelli VRT, Bodenmiller B. ImcSegmentationPipeline: A pixelclassification based multiplexed image segmentation pipeline. In. Zenodo2017.
42. Schapiro D, et al. histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nature methods.* 2017;14(9):873-876.
43. Van Gassen S, et al. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A.* 2015;87(7):636-645.
44. Levine JH, et al. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell.* 2015;162(1):184-197.
45. Amir el AD, et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat Biotechnol.* 2013;31(6):545-552.
46. Borg IG, P. . *Modern Multidimensional Scaling: theory and applications (2nd ed.).* . New York: Springer-Verlag; 2005.
47. Wisniewski JR. Proteomic sample preparation from formalin fixed and paraffin embedded tissue. *J Vis Exp.* 2013(79).
48. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Molecular & cellular proteomics : MCP.* 2014;13(9):2513-2526.
49. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med.* 1990;9(7):811-818.
50. Huber W, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nature methods.* 2015;12(2):115-121.
51. Illumina. Illumina Technical Note: Gene Expression Microarray Data Quality Control. 2007.
52. Curtis C, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature.* 2012;486(7403):346-352.
53. Comprehensive molecular portraits of human breast tumours. *Nature.* 2012;490(7418):61-70.

54. Parker JS, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* 2009;27(8):1160-1167.

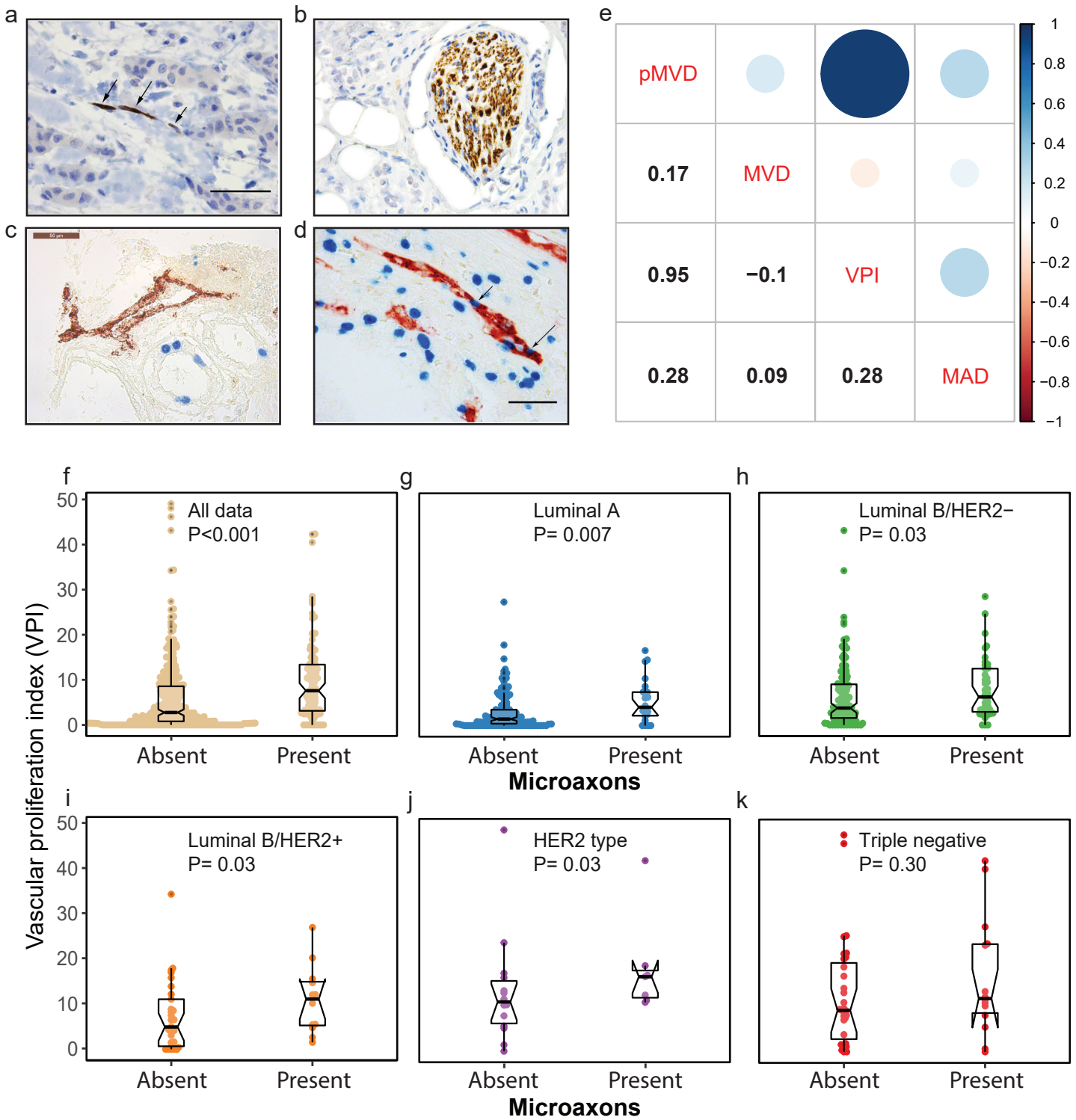

Figure 1. Immunostaining for markers of angiogenesis and neurogenesis and their associations in primary breast cancer (Bergen Breast Cancer cohort, n=461-483). **a**) Immunostaining with Neurofilament antibody (x400), illustrating the positivity in the microaxons between the cords and glands of tumor tissue; **b**) Presence of thick Neurofilament positive nerve bundles; **c-d**) Dual immunostaining with Factor VIII (red) for endothelial cells and Ki67 (blue) for proliferating cells (x400). Vessels with no proliferative activity in (**c**), and Ki67 positive proliferating endothelial cells in (**d**) (highlighted with arrows); **e**) Correlation plot showing pairwise correlation between proliferating microvessel density (pMVD), microvessel density (MVD), vascular proliferation index (VPI), and microaxon density (MAD). The lower triangular part shows the Spearman's correlation coefficient, the upper triangular part shows the same correlation as areas of circles; **f-k**) Boxplots showing the Vascular proliferation index (VPI) across breast cancer cases with stromal micoraxons absent or present (primary tumors; All P values were based on Mann-Whitney U test): (**f**) All primary tumors (n=461), (**g**) Luminal A (n=179), (**h**) Luminal B/HER2 negative (n=171), (**i**) Luminal B/HER2 positive (n=44), (**j**) HER2 positive (n=21) and (**k**) Triple negative (n=46).

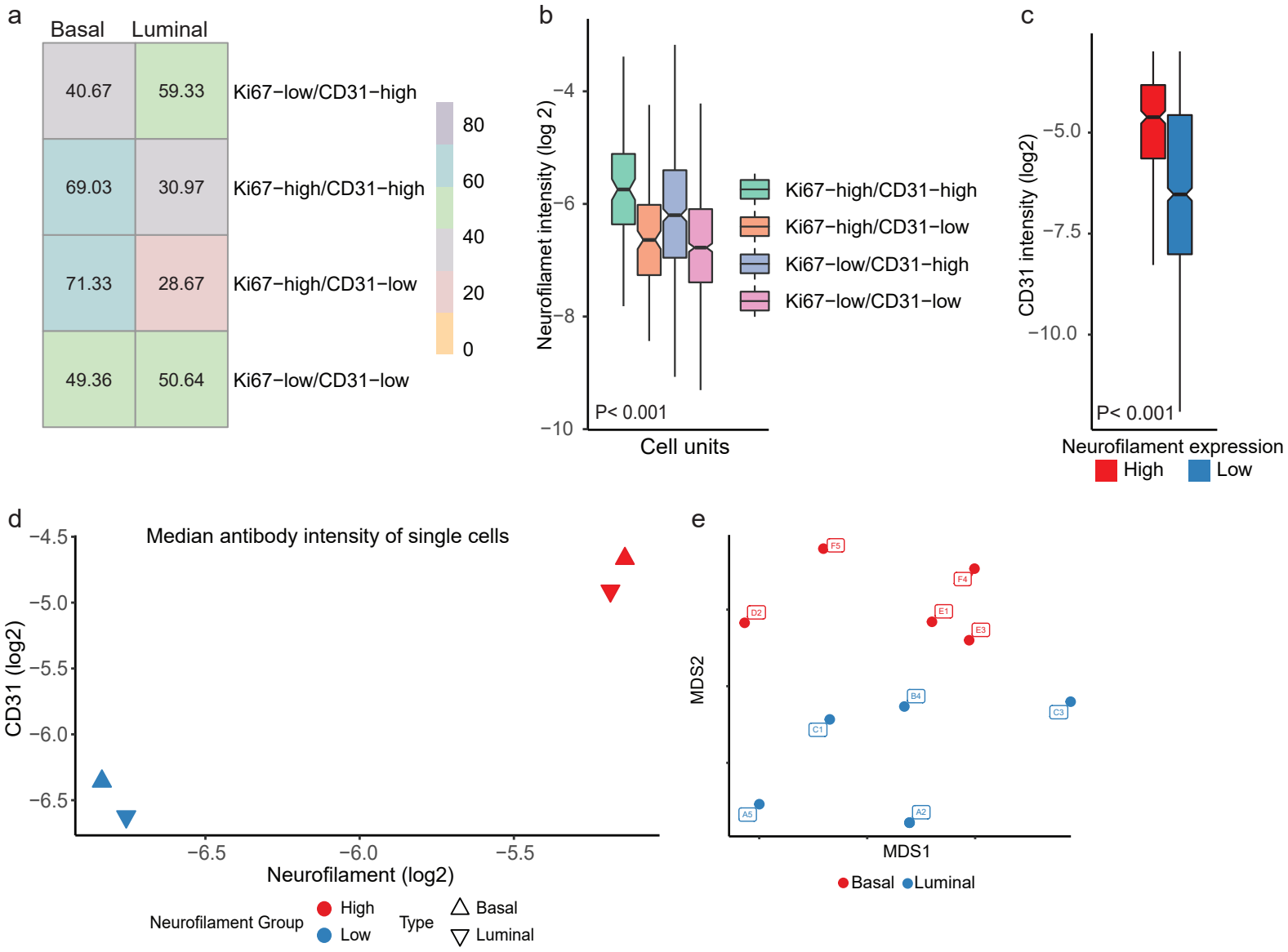

Figure 2. Investigating features of neurogenesis and angiogenesis in breast cancer using Imaging Mass Cytometry (IMC) data. **a**) Heatmap showing proportion (%) of single-cells from basal-like and luminal-like IMC samples. The endothelial cells were gated into four subgroups namely Ki67-low/CD31-high, Ki67-high/CD31-high, Ki67-high/CD31-low, and Ki67-low/CD31-low, on the basis of their Ki67 and CD31 antibody intensities; **b**) Boxplots showing the Neurofilament antibody intensity (log2 transformed) of all endothelial single-cells, stratified into four Ki67/CD31 subgroups from (**a**); **c**) Boxplot showing CD31 antibody intensity (log2 transformed) of single-cells stratified to high and low subgroups of Neurofilament antibody intensity; **d**) Scatter plot showing the median intensity values of CD31 and Neurofilament antibodies. Data points were annotated on the basis of their molecular subtype (basal-like vs. luminal-like), and their corresponding Neurofilament intensity subgroup (high vs. low); **e**) Multidimensional scaling plot of all basal-like and luminal-like IMC (n=10) samples included in the study. The plot visualizes the 2D mapping of samples based on an initial set of features describing neural and vascular characteristics.

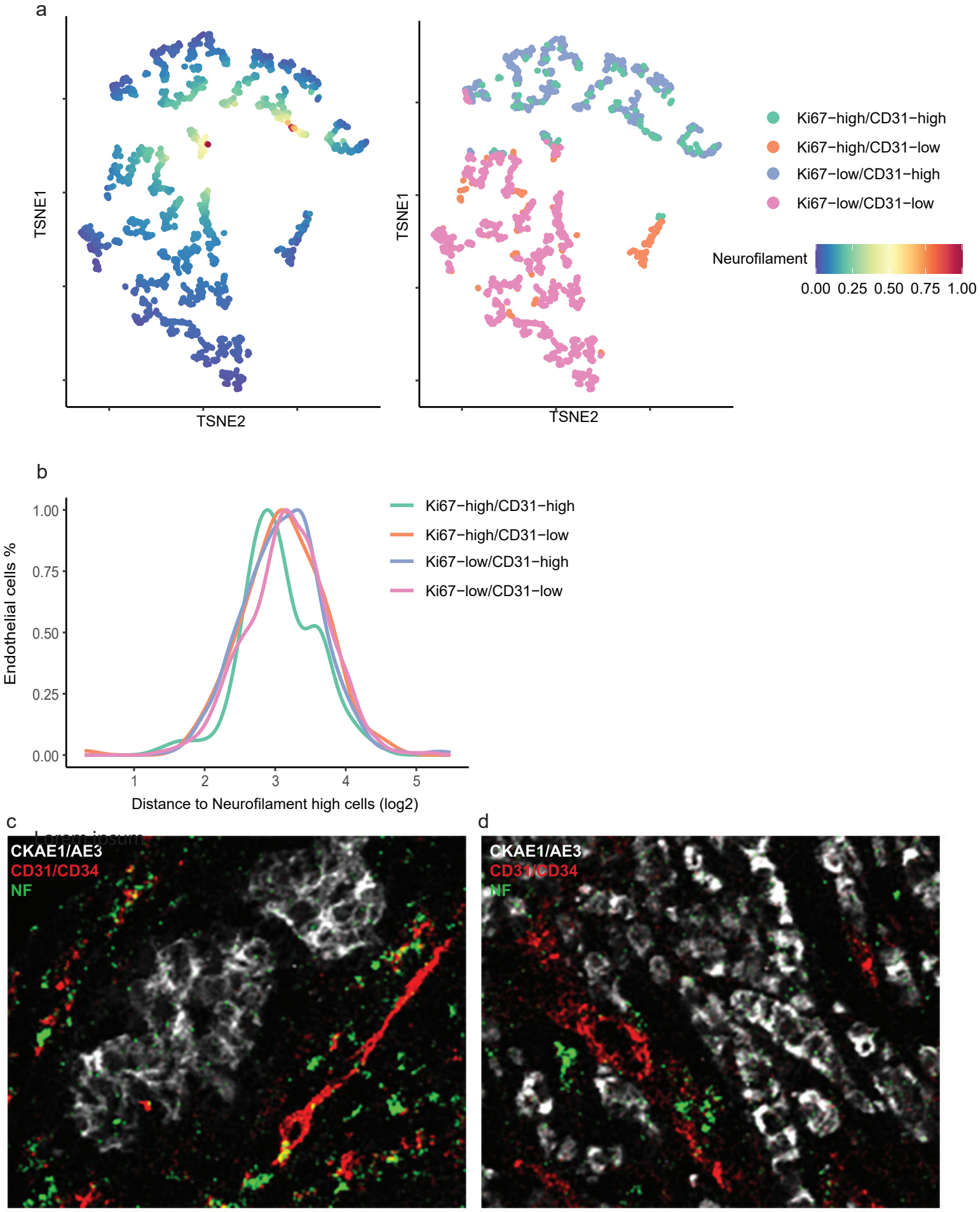

Figure 3. Co-expression and spatial co-localization of neural and vascular markers and structures. **a**) 2D map generated using TSNE of all endothelial single cells (n=1,868) from 10 IMC breast cancer tumors. The left-most map is colored using the Neurofilament antibody intensity levels of single cells. The right-most subpanel shows the same 2D map, annotated based on the identified endothelial subgroups; **b**) Log2-transformed average distances of endothelial cells to their neighboring cells of Neurofilament high expression. The distance estimation is based on k=5 nearest neighbors from the endothelial cells of interest; **c-d**) Representative IMC images showing spatial co-localization of neural and vascular structures. The images are pseudo colored using the antibody intensity values of Neurofilament (green), CD31/34 (red) and the cytokeratin CKAE1/AE3 (white/grey). Focal co-expression of Neurofilament and CD31/CD34 is observed (yellow).

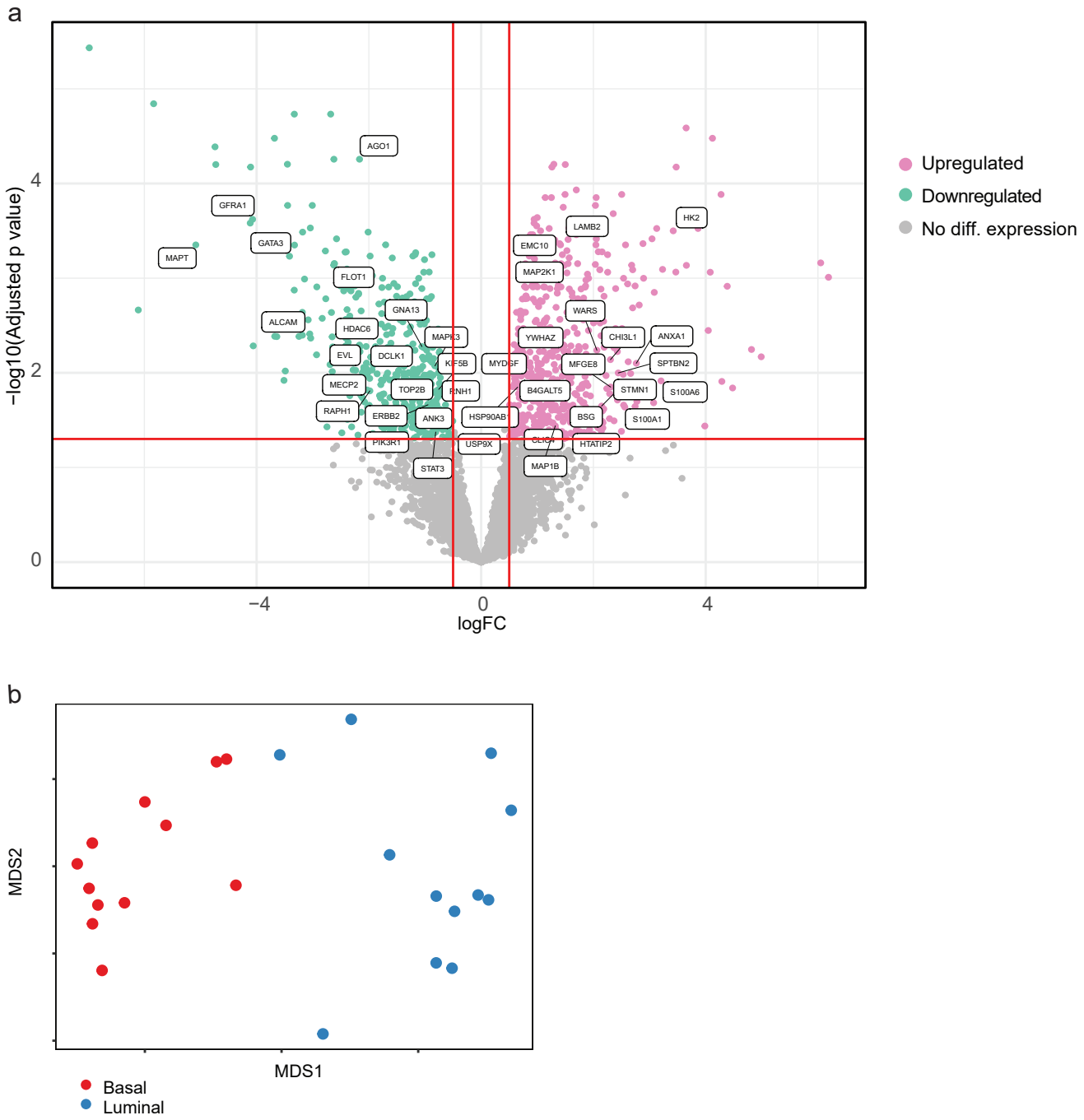

Figure 4. Analysis of neurogenesis-realted and angiogenesis-related proteins in breast cancer tissues using mass spectrometry. **a**) Volcano plot showing differentially enriched proteins between basal-like and luminal-like tissues. P-values were adjusted using Benjamin-Hochberg method. **b)** Multidimensional scaling plot of basal-like and luminal-like tumors using mass spectrometry. Scatter plot showing the 2D mapping of all samples (n=24). The initial dimensions were the protein expression levels of all neurogenic and angiogenic proteins that were differentially regulated from figure **a**.

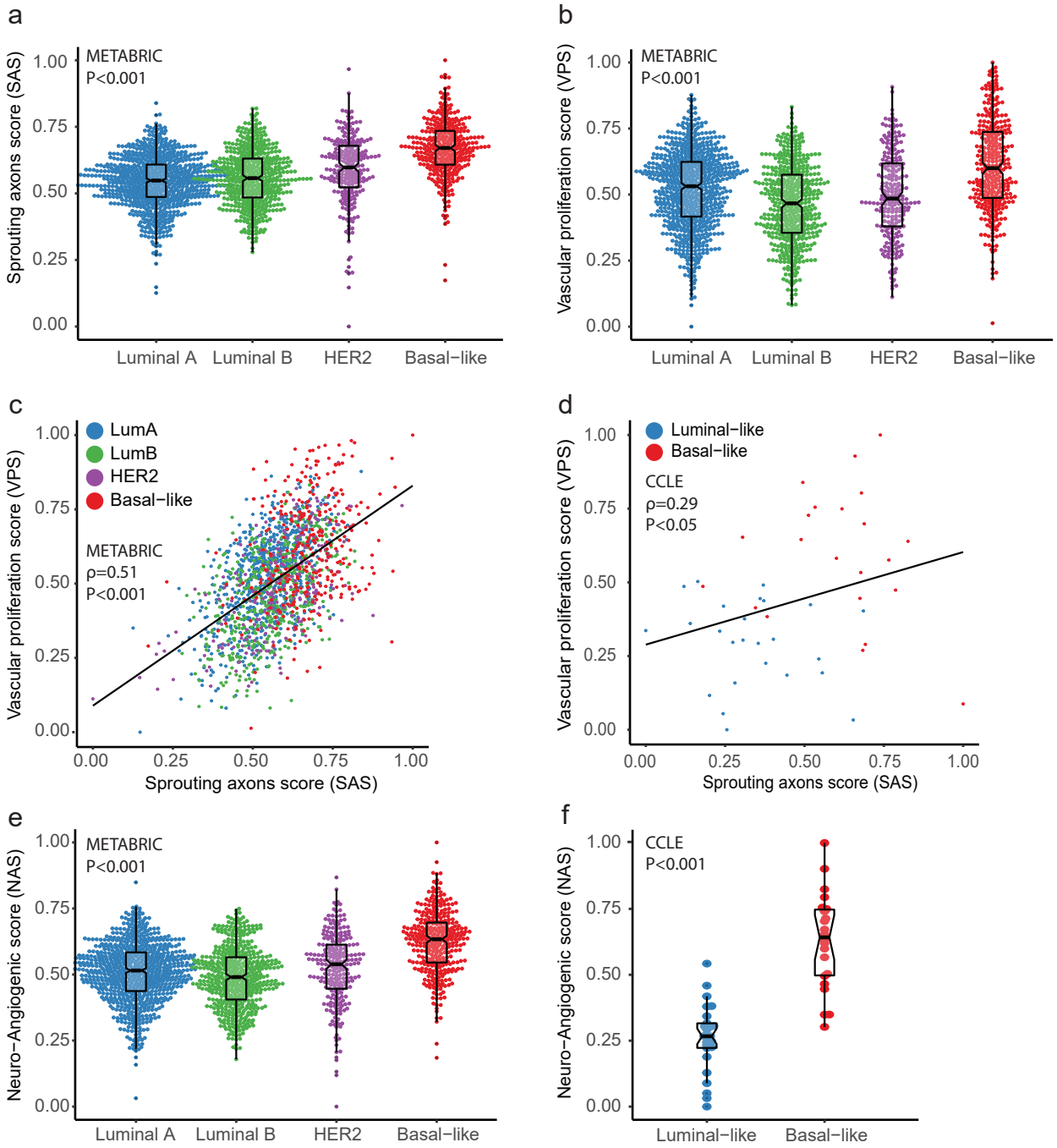

Figure 5. Gene expression analysis of neurogenic and angiogenic features across molecular breast cancer subtypes, METABRIC cohort, n=1791; Cancer Cell Line Encyclopedia (CCLE), Breast Cancer cell lines, n=47. **a**) Bee swarm plots showing the distribution of Sprouting axons score (SAS) across molecular subtypes of breast cancer; **b**) Bee swarm plots showing the distribution of Vascular proliferation score (VPS) across molecular subtypes of breast cancer; **c**) Scatter plot showing the relationship between SAS and VPS, using data from the METABRIC cohort; **d**) Scatter plot showing the relationship between SAS and VPS, using data from all CCLE breast cancer cell lines; **e)** Boxplots showing the distribution of Neuro-angiogenic score (NAS) across molecular subtypes using data from the METABRIC cohort; **f**) Boxplots showing the distribution of NAS across molecular subtypes using data from CCLE breast cancer cell lines. Mann-Whitney U or Kruskal-Wallis tests were used for comparing continuous variables in different groups. Correlations were tested by Spearman's rank correlation test, and the Spearman's correlation coefficient (ρ) is reported. Molecular subtypes are indicated with different colors. In METABRIC data: Luminal A - blue; Luminal B – green; HER2 enriched – purple; Basal-like - red. In the CCLE data: Luminal-like – blue; Basal-like – red.

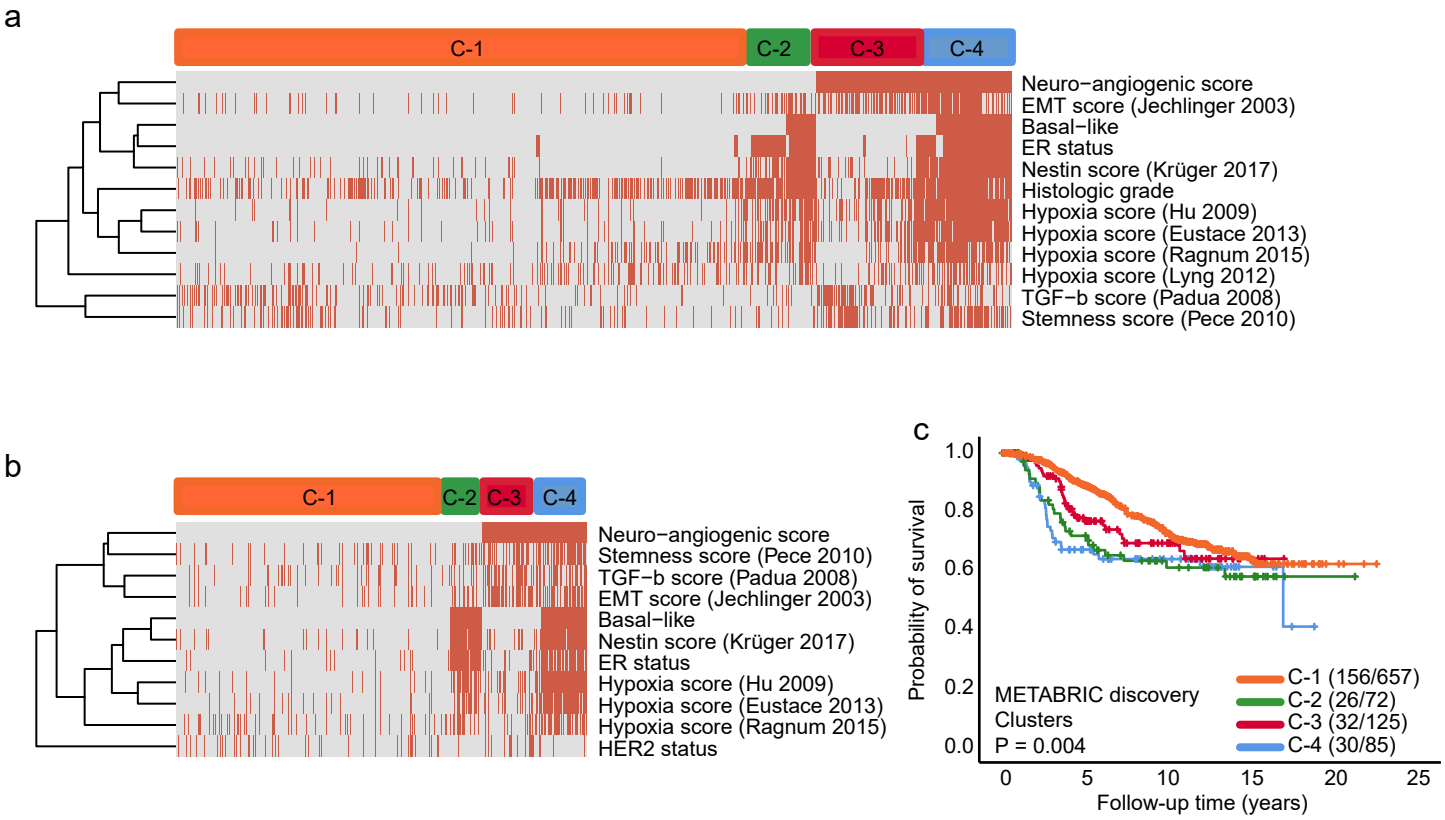

Figure 6. Exploratory analysis of heterogeneous phenotypes among breast cancers with high/low Neuro-angiogenic score. **a-b**) Heatmap showing the results of hierarchical clustering analysis conducted on patients from: **a**) the METABRIC discovery cohort (n=939), and **b**) the TCGA cohort (n=505). Clustering was conducted on binary variables reflecting either high (red color) / low (gray color) signature scores, positivity (red color) / negativity (gray color) for ER and HER2, or presence (red color) / absence (gray color) of histologic grade 3, and basal-like phenotype from PAM50. Four cluster groups were evident visually, and the tree was cut horizontally to further explore these clusters marked as C1, C2, C3, C4; **c**) Kaplan-Meier curve illustrating the difference in survival between patients in the cluster groups C1-C4 using data from the METABRIC discovery cohort from (**a**). Differences in breast cancer specific survival between groups were tested using the log-rank test.

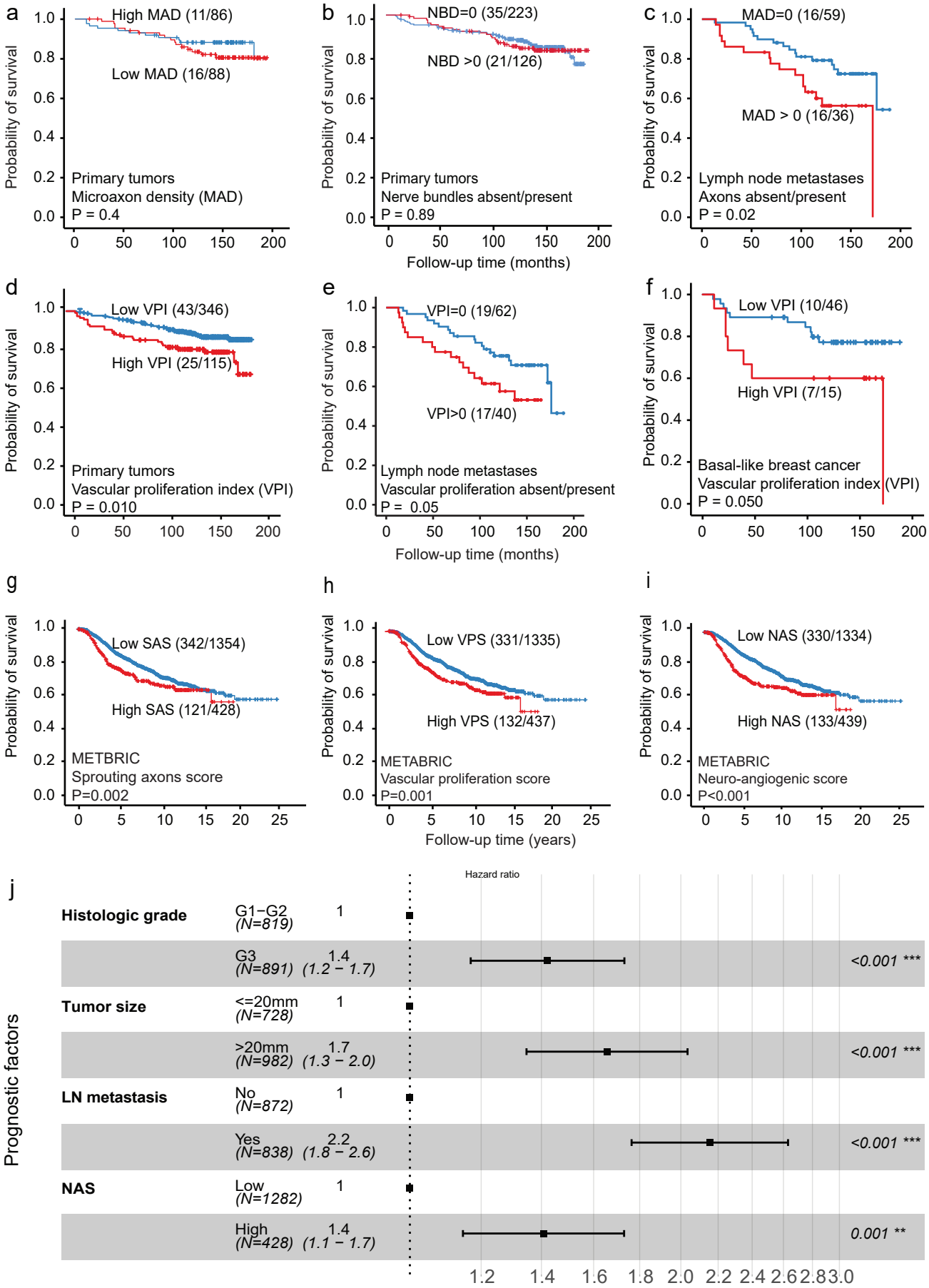

Figure 7. Markers of angiogenesis, neurogenesis, and neuro-angiogenesis and their association with survival in breast cancer. **a-b**) Microaxon density (MAD), and nerve bundle density (NBD) do not associate with survival in primary breast cancers; **c**) Presence of microaxons in lymph node (LN) metastases associates with reduced breast cancer specific survival; **d-f**) High vascular proliferation index (VPI) in primary breast cancer, and LN metastases associates with reduced cancer specific survival, also in the subset of primary basal-like breast cancers; **g-h**) High Sprouting axons score (SAS), Vascular proliferation score (VPS), and Neuro-angiogenic score (NAS) associate with reduced breast cancer specific survival in patients from the METABRIC cohort. In all Kaplan-Meier analyses the log-rank test was used to quantify differences. Next to every survival curve we report the number of breast cancer deaths, out of the total number of cases in the category; **j**) Cox' proportional hazard modelling shows that the Neuro-angiogenic score demonstrates independent prognostic value in multivariate survival analysis, adjusting for tumor size, histologic grade, and LN status, using data from the combined METABRIC cohort.

Table 1. Associations between presence of axons and clinico-pathologic features of primary tumors (n=483)

Variable	Axons (primary tumor)				
	Absent n (%)	Present n (%)	OR	(95% CI)	P[C]
Tumor diameter					0.001
≤ 2 cm	302 (82)	65 (18)	1.0		
>2 cm	79 (68)	37 (32)	2.2	(1.4-3.5)	
Histologic grade					<0.001
Grade 1 - 2	325 (82)	71 (18)	1.0		
Grade 3	56 (64)	31 (36)	2.5	(1.5-4.2)	
Lymph node metastasis					NS
No	281 (81)	68 (19)	1.0		
Yes	96 (74)	33 (26)	1.4	(0.9-2.3)	
ER					0.027
Positive	326 (81)	78 (19)	1.0		
Negative	55 (70)	24 (30)	1.8	(1.1-3.1)	
PR					NS
Positive	270 (81)	65 (19)	1.0		
Negative	111 (75)	37 (25)	1.4	(0.9-2.2)	
HER2 status					0.027
Negative	335 (81)	81 (19)	1.0		
Positive	46 (69)	21 (31)	1.9	(1.1-3.3)	
Ki-67 %[A]					<0.001
Low	291 (85)	53 (15)	1.0		
High	90 (65)	49 (35)	3.0	(1.9-4.7)	
Molecular subtype[B]					
Luminal A	165 (88)	23 (12)			
Luminal B / HER2 -	138 (76)	44 (24)			0.003
Luminal B / HER2+	30 (70)	13 (30)			0.003
HER2 +	16 (67)	8 (33)			0.006
Triple negative	32 (70)	14 (30)			0.002
Basal-like phenotype (CK5/6)					0.09
Negative	336 (80)	84 (20)	1		
Positive	45 (71)	18 (29)	1.6	(0.9-2.9)	

n: number of patients; OR: odds ratio; CI: confidence interval; ER: Estrogen receptor; PR: Progesterone receptor; HER2: Human epidermal growth factor 2; CK5/6: Cytokeratin 5/6. [A]Cut-off 30% (upper quartile); [B] St. Gallen 2013; [C]P-values by Pearson's chi-square test. Missing data on lymph node status, n = 5.

Table 2. Associations between markers of vascular proliferation (pMVD, VPI) and clinico-pathologic features of primary tumors (n= 461).

Variable	n (%)	pMVD[D]	P[E]	VPI[D]	P[E]
Tumor diameter			0.007		0.002
≤ 2 cm	344 (75)	2.6		3.07	
>2 cm	117 (25)	4.3		5.97	
Histologic grade			<0.001		<0.001
Grade 1 - 2	378 (82)	2.2		2.78	
Grade 3	83 (18)	7.0		9.62	
Lymph node metastasis[A]			0.033		0.023
No	331 (72)	2.6		3.15	
Yes	127 (28)	4.3		4.13	
ER			<0.001		<0.001
Positive	385 (83)	2.1		2.8	
Negative	76 (17)	7.8		10.8	
PR			<0.001		<0.001
Positive	322 (70)	2.1		2.80	
Negative	139 (30)	5.0		7.47	
HER2 status			0.001		<0.001
Negative	396 (86)	2.6		2.90	
Positive	65 (14)	6.1		8.88	
Ki-67 %[B]			<0.001		<0.001
Low	328 (71)	1.7		2.29	
High	133 (29)	7.4		9.62	
Molecular subtype[C]			<0.001		<0.001
Luminal A	179 (39)	1.3		1.56	
Luminal B/ HER2 -	171 (37)	3.5		4.03	
Luminal B / HER2+	44 (9)	5.0		6.01	
HER2 +	21 (5)	7.0		12.4	
Triple negative	46 (10)	8.5		10.5	
Basal-like phenotype (CK5/6)			<0.001		<0.001
Negative	399 (87)	2.6		3.02	
Positive	62 (13)	6.7		8.83	

n: number of cases; MVD: microvessel density; pMVD: proliferating vascular density; VPI: vascular proliferation index; ER: estrogen receptor; PR: Progesterone receptor; HER2: Human epidermal growth factor 2. [A] Missing information on lymph node metastasis, *n*=3; [B] Cut off: 30% (upper quartile); [C] St.Gallen 2013; [D] Continuous variables (median value reported for each category); [E] Mann-Whitney U test.