EXPRESSION OF CHEMOKINE RECEPTORS CXCR4 AND CXCR7 IN CIRCULATING TUMOR CELLS OF BREAST CANCER

Grigoryeva ES[1]✉, Tashireva LA[1], Savelieva OE[2], Zavyalova MV[1], Cherdyntseva NV[1], Perelmuter VM[1]

[1] Cancer Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia

[2] Saint Petersburg State Pediatric Medical University, Saint Petersburg, Russia

It is known that chemokine receptors CXCR4 and CXCR7 in primary tumor cells are associated with tumor growth progression; however, the significance of their expression in circulating tumor cells (CTCs) remains insufficiently studied. The objective of this study was to investigate the expression of chemokine receptors CXCR4 and CXCR7 in subpopulations of CTCs with positive (EpCAM$^+$) and negative (EpCAM$^-$) EpCAM expression in breast cancer patients, as well as assessed their correlation with clinicopathological parameters and prognostic relevance. The study methods included protein expression analysis and transcriptome profiling of CTCs obtained from peripheral blood. This study comprehensively characterized CXCR4 and CXCR7 expression across EpCAM$^+$ and EpCAM$^-$ CTC subsets and assessed their clinical relevance through protein-level detection, transcriptomic profiling, and long-term patient follow-up. CXCR4 and CXCR7 receptors were predominantly expressed in EpCAM$^+$ CTCs, whereas EpCAM$^-$ cells were largely negative. Importantly, an increased number of EpCAM$^-$ CTCs, irrespective of CXCR4/7 status, was associated with disease progression over a six-year period ($p = 0,0007$). Prognostic significance was specifically attributed to EpCAM$^-$CXCR4/7$^-$ CTCs, with counts exceeding 1.25 cells/ml predicting progression with high sensitivity and specificity. Distinct CTC subpopulations were further characterized by stemness and epithelial-mesenchymal transition (EMT) markers, underscoring the aggressive phenotype of EpCAM$^-$ cells exhibiting EMT traits. Transcriptomic analysis of *EpCAM$^-$CXCR4/7$^-$* CTCs revealed upregulation of genes involved in ferroptosis ($p = 3.315 \times 10^{-7}$) and androgen receptor signaling pathways ($p = 8.0$ v 10^{-5}), alongside identification of progression-associated genes (*HBB*, *IGLC2*, and *IGHM*). Conversely, *MALAT1* was overexpressed in patients without progression, indicating a potential metastasis-suppressive function ($p = 1.52 \times 10^{-2}$). These findings highlight the pathogenetic importance of EpCAM$^-$ CTCs in breast cancer progression and support a paradigm shift in CTC research towards this subpopulation. Further investigations are warranted to elucidate the functional roles of these cells and their utility as prognostic biomarkers.

Keywords: circulating tumor cells, chemokine receptors, epithelial-mesenchymal transition, stemness, transcriptomics, breast cancer progression

ЭКСПРЕССИЯ РЕЦЕПТОРОВ ХЕМОКИНОВ CXCR4 И CXCR7 В ЦИРКУЛИРУЮЩИХ ОПУХОЛЕВЫХ КЛЕТКАХ ПРИ РАКЕ МОЛОЧНОЙ ЖЕЛЕЗЫ

Е. С. Григорьева[1]✉, Л. А. Таширева[1], О. Е. Савельева[2], М. В. Завьялова[1], Н. В. Чердынцева[1], В. М. Перельмутер[1]

[1] Научно-исследовательский институт онкологии, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия

[2] Санкт-Петербургский государственный педиатрический медицинский университет, Санкт-Петербург, Россия

Хемокиновые рецепторы CXCR4 и CXCR7 в клетках первичной опухоли связаны с прогрессией опухолевого роста. Цель исследования — изучить экспрессию CXCR4 и CXCR7 в субпопуляциях ЦОК с положительной (EpCAM$^+$) и отрицательной (EpCAM$^-$) экспрессией EpCAM у пациентов с раком молочной железы (РМЖ), а также оценить их связь с клинико-патологическими параметрами и прогностической значимостью. Анализ белковой экспрессии и транскриптомное профилирование ЦОК, полученных из периферической крови, показали, что CXCR4 и CXCR7 преимущественно экспрессировались в EpCAM$^+$ ЦОК, тогда как EpCAM$^-$ клетки в основном не имели экспрессии CXCR4 и CXCR7 ($p < 0,001$). Важно, что увеличение количества EpCAM$^-$ ЦОК, независимо от экспрессии CXCR4/7, было связано с прогрессированием заболевания в течение шестилетнего периода ($p = 0,0007$). Прогностическое значение показано для EpCAM$^-$CXCR4/7$^-$ ЦОК, при этом количество, превышающее 1,25 кл./мл периферической крови, предсказывало прогрессирование заболевания с чувствительностью 100% и специфичностью 95,1%. Отдельные субпопуляции ЦОК были дополнительно охарактеризованы по маркерам стволовости и эпителиально-мезенхимального перехода (ЭМП), что подчеркивает агрессивный фенотип EpCAM$^-$ клеток, проявляющих признаки ЭМП. Транскриптомный анализ *EpCAM$^-$CXCR4/7$^-$* ЦОК выявил повышенную экспрессию генов, участвующих в ферроптозе ($p = 3,315 \times 10^{-7}$) и сигнальных путях андрогенового рецептора ($p = 8,0 \times 10^{-5}$), а также генов, ассоциированных с прогрессированием (*HBB*, *IGLC2* и *IGHM*). Напротив, *MALAT1* был сверхэкспрессирован у пациентов без прогрессирования, что указывает на потенциальную подавляющую метастазирование функцию ($p = 1,52 \times 10^{-2}$). Результаты подчеркивают патогенетическую значимость EpCAM$^-$ ЦОК в прогрессировании РМЖ и поддерживают смену парадигмы в исследованиях ЦОК в сторону этой субпопуляции.

Ключевые слова: циркулирующие опухолевые клетки, хемокиновые рецепторы, эпителиально-мезенхимальный переход, стволовые клетки, транскриптомика, прогрессирование рака молочной железы

Breast cancer (BC) remains one of the most significant challenges in modern oncology due to its high prevalence, substantial mortality rate, and frequent metastasis. The recent advances in liquid biopsy techniques have opened new avenues for developing approaches to better predict the clinical course of the disease. Among the most promising targets for investigation are circulating tumor cells (CTCs), which are tumor cells that have detached from the primary tumor and circulate in the bloodstream. Elevated levels of CTCs in the peripheral blood often correlate with poor prognosis, making them a valuable tool for patient stratification and optimizing treatment strategies [1]. However, the clinical application of CTCs faces several challenges, including their low concentration in blood and technical difficulties in their detection [2]. Most studies on CTCs rely on EpCAM (epithelial cell adhesion molecule)-based methods [3]. This preference is largely due to the availability of commercially certified systems designed to detect EpCAM-positive CTCs [4]. Although EpCAM-expressing tumor cells have traditionally been considered the primary drivers of metastasis, increasing evidence highlights the important role of subpopulations lacking EpCAM expression [5]. The loss of membrane EpCAM expression is often associated with epithelial-mesenchymal transition (EMT), a process in which epithelial cells lose polarity and cell-cell adhesion, acquiring mesenchymal traits such as enhanced motility and resistance to apoptosis [6]. In tumor progression, EMT facilitates tumor cell motility, invasion, intravasation, survival in circulation, and the formation of metastases following extravasation into distant organs [7]. Consequently, CTCs constitute a heterogeneous population comprising cells with varying epithelial and mesenchymal phenotypes, as well as stem cell-like properties, which may underlie their metastatic potential.

In the context of investigating CTC heterogeneity, the chemokine receptors CXCR4 and CXCR7, which are activated by their common ligand SDF-1 (CXCL12), are of considerable interest. The most well-characterized function of the CXCR4–SDF-1 axis is to mediate the directed migration of bone marrow progenitor cells and immune cells to sites of inflammation. It is well established that CXCR4 and CXCR7 are frequently overexpressed in tumor cells, and their interaction with SDF-1 plays a critical role in cancer progression and metastasis [8]. Tumor cells with elevated CXCR4 expression exhibit increased proliferation, driven by activation of the MAPK and PI3K/Akt signaling pathways, while enhanced levels of anti-apoptotic proteins and reduced expression of death receptors promote cell survival [9]. Initially, CXCR7 was considered a decoy receptor that sequesters CXCL12, thereby attenuating CXCR4 activity [10]. However, recent evidence reveals that CXCR7 can signal via the noncanonical β-arrestin pathway, leading to activation of intracellular cascades including protein kinase B (AKT) and JAK/STAT pathways, which further stimulate tumor cell proliferation and migration [11]. According to the literature, high CXCR4 expression correlates with an increased risk of breast cancer metastasis to lymph nodes and distant organs, as well as with reduced relapse-free and overall survival [12]. In summary, while the expression patterns of CXCR4 and CXCR7 in primary breast tumors have been extensively characterized, studies examining their expression in circulating tumor cells remain very limited.

Thus, the aim of this study is to investigate the expression of chemokine receptors CXCR4 and CXCR7 in CTC subpopulations with positive (EpCAM+) and negative (EpCAM−) EpCAM expression in breast cancer patients, and to evaluate their association with clinicopathological parameters and prognostic significance.

METHODS

Patients

The study included 65 female patients with invasive breast carcinoma of no special type, who were treated at the Oncology Research Institute Clinic of the Tomsk National Research Medical Center (Tables 1, 2). CTC analysis was performed prior to any treatment. Patients received full treatment according to the clinical guidelines of the Ministry of Health of the Russian Federation. The follow-up period was 6 years.

Flow cytometry

Venous blood samples collected from breast cancer patients was used for CTC detection. Cell concentrates were prepared by sedimentation, followed by collection of the white cell layer at the interface between the erythrocyte sediment and the separated plasma, as well as the entire supernatant, according to the method described by R. A. Pospelova [13].

Table 1. Clinicopathological characteristics of patients included in the study of CTC subpopulation composition

Parameter		Frequency, % (*n*)
Age	≤ 50	26.67% (12/45)
	> 50	73.33% (33/45)
Tumor size (cT)	1	35.56% (16/45)
	2	60.00% (27/45)
	4	4.44% (2/45)
Stage	IA	28.89% (13/45)
	IIA	33.33% (15/45)
	IIB	33.33% (15/45)
	IIIB	4.44% (2/45)
Lymph node status (cN)	0	53.33% (24/45)
	1	46.67% (21/45)
Molecular subtype	Luminal A	26.67% (12/45)
	Luminal B (HER2−)	37.78% (17/45)
	Luminal B (HER2+)	20.00% (9/45)
	Triple negative	6.67% (3/45)

Table 2. Clinicopathological characteristics of patients included in the study of the transcriptional profile of circulating tumor cells

Parameter		Frequency, % (n)
Age	≤ 50	30.0% (6/20)
	> 50	70.0% (14/20)
Tumor size (cT)	1	5.0% (1/40)
	2	80.0% (15/40)
	3	5.0% (1/20)
	4	5.0% (3/20)
Stage	I	5.0% (1/20)
	IIA	20.0% (5/20)
	IIB	40.0% (8/20)
	IIIA	15.0% (3/20)
	IIIB	10.0% (2/20)
	IIIC	5.0% (1/20)
Lymph node status (cN)	0	30.0% (6/20)
	1	70.0% (14/20)
Molecular subtype	Luminal A	10.0% (2/20)
	Luminal B (HER2⁻)	65.0% (13/20)
	Luminal B (HER2⁺)	10.0% (2/20)
	Triple negative	5.0% (5/20)
	HER2⁺	10.0% (1/20)

Samples for flow cytometry were prepared as follows. The cell concentrate was washed by adding 1 ml of CellWASH solution (BD Biosciences, USA) and centrifuged at 300 × g for 10 minutes. To lyse erythrocytes, 500 μl of OptiLyse C buffer (Beckman Coulter, France) was added, and the samples were washed with 2 ml of CellWASH solution for 10 minutes at 300 × g, followed by removal of the supernatant. After blocking nonspecific Fc receptor binding with Human TruStain FcX™ Fc Receptor Blocking Solution (BioLegend, USA), 5 μl of the following monoclonal antibodies were added to the cell concentrate: BV570 anti-human CD45 (clone HI30; Sony Biotechnology, USA), BV650 anti-human CD326 (EpCAM) (clone 9C4; Sony Biotechnology, USA), BV510 anti-human CD44 (clone G44-26; BD Horizon, USA), PerCP/Cy5.5 anti-human CD24 (clone ML5; Sony Biotechnology, USA), BV421 anti-human CXCR4 (clone 12G5; Sony Biotechnology, USA), BV421 anti-human CXCR7 (clone 10D1; BD Biosciences; USA) and PE/Cy7 anti-human N-Cadherin (clone 8C11; Sony Biotechnology, USA). The samples were then incubated in the dark at room temperature for 20 minutes.

Following incubation, probes were washed with 2 ml of CellWASH solution for 10 minutes at 300 × g and removing the supernatant. For intracellular staining, 250 μl of BD Cytofix/Cytoperm solution (BD Biosciences, USA) was added to each unstained and stained sample, followed by incubation in the dark at 4 °C for 20 minutes. Samples were then washed twice in 1 ml of BD Perm/Wash buffer (BD Biosciences, USA) by centrifugation at 300 × g for 6 minutes. Subsequently, 50 μl of BD Perm/Wash buffer was added to each sample, along with 5 μl of the following antibodies: AF647-anti-human CK7/8 (clone CAM5.2; BD Pharmingen, USA), AF488-anti-human Snail1 (clone 20C8; eBioscience, USA), and AF750-anti-human Vimentin (R&D Systems, USA). The samples were incubated at 4 °C for 20 minutes.

Each sample was then washed in 1 ml of CellWASH buffer (BD Biosciences; USA) by centrifugation at 300 g for 6 min. At the final stage, 100 μl of Cell Staining Buffer (Sony Biotechnology; USA) were added to the sediment and the sample was resuspended.

Samples were analyzed on a Novocyte 3000 flow cytometer (ACEA Biosciences; USA) using NovoExpress 1.3.0 (ACEA Biosciences; USA). The concentration of circulating cells was calculated per 1 ml of blood.

Statistical analysis was performed using the Prism 10.4.1 package (GraphPad Software; USA). The Kruskal–Wallis test was used to analyze independent groups, and the Mann–Whitney test was used to analyze two independent groups. ROC analysis was used to assess the prognostic value of the prediction accuracy. Differences were considered significant at $p < 0.05$.

scRNA-seq analysis of CTCs

Public scRNA-seq data set from 20 BC patients (T1-4N0-3M0, all molecular subtypes) generated in our previous study [14] and available via BioProject under the accession number PRJNA776403 was used for investigation of transcriptional profile of EPCAM-negative and EPCAM-positive CTCs.

The Seurat software package, version 4.0.4 [15], was employed for quality control and analysis of single-cell RNA sequencing data. Cell doublets were identified using DoubletFinder [16] and subsequently removed from each dataset. Integration of the 20 datasets with default parameters was performed. The aggregated data underwent preprocessing, involving the exclusion of cells with unique feature counts less than 200 and mitochondrial percent exceeding 25. Raw RNA UMI counts of the aggregated data were normalized, followed by principal component analysis (PCA). The dataset was visualized and explored using the uniform manifold approximation and projection (UMAP) method, a nonlinear dimensional reduction technique.

Spatial transcriptomics data analysis of breast tumor tissue

Spatial transcriptomics dataset generated in our previous study [17] and available via GEO Database under the accession number GSE242311 was used to investigate gene expression of EPCAM-negative and EPCAM-positive tumor cells in primary tumor of five BC patients (invasive carcinoma of nonspecific type, luminal A and B, stage I–IIA, grade 2–3). Samples
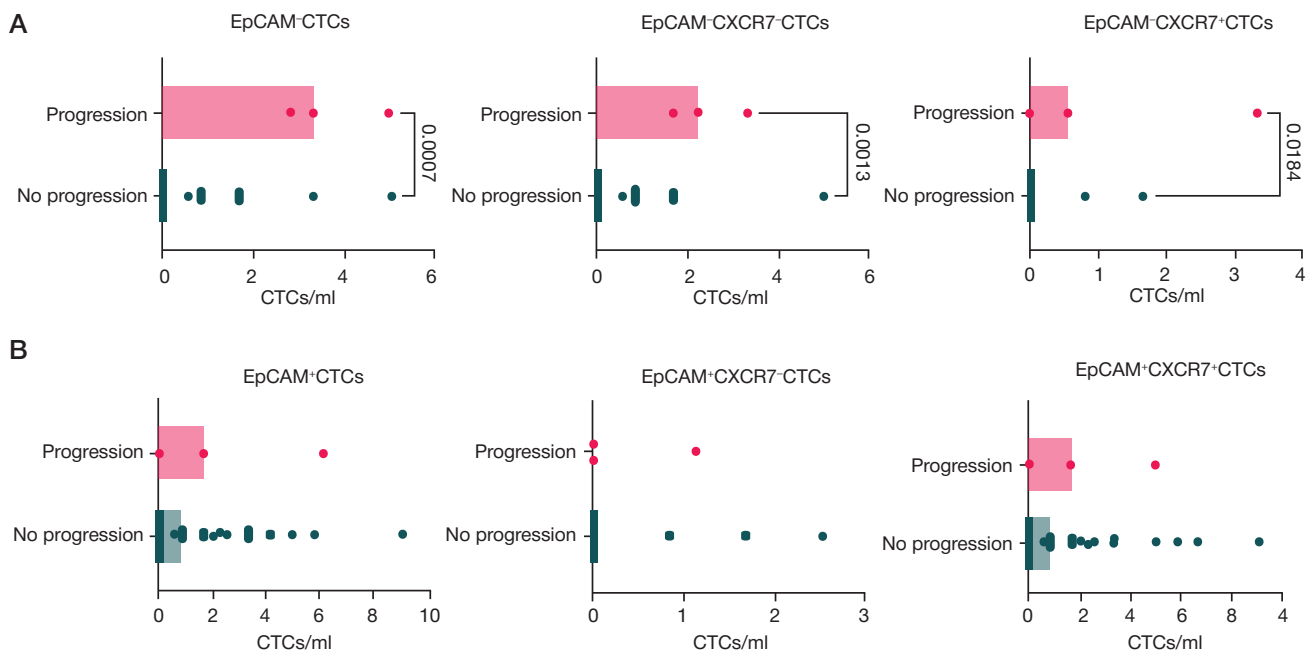
Fig. 1. Number of EpCAM⁻ (**A**) and EpCAM⁺ (**B**) CTCs according to CXCR4/7 expression in breast cancer patients with tumor progression during the 6-year follow-up period

were filtered, excluding genes with nonzero expression in fewer than 10 tissue spots and tissue spots with fewer than 200 filtered genes. The raw counts were normalized using the SCTransform [18] function with default parameters. The uniform manifold approximation and projection (UMAP) technique were then applied to the SCTransform-normalized counts, utilizing the first 30 principal components determined through principal component analysis (PCA). The results were visualized using the Seurat package.

RESULTS

Association between the Number of EpCAM⁺ and EpCAM⁻ CTCs Expressing CXCR4 and CXCR7 and Clinical Parameters in BC Patients

The number of EpCAM⁺ and EpCAM⁻ CTCs was evaluated in the peripheral blood of breast cancer patients. EpCAM⁺ CTCs were defined as cells expressing EpCAM but lacking the leukocyte common antigen CD45, irrespective of cytokeratin 7/8 expression. In contrast, EpCAM⁻ CTCs were defined as CD45-negative cells without EpCAM expression but positive for cytokeratins 7/8. The analysis revealed that the number of EpCAM⁺ CTCs was significantly higher than that of EpCAM⁻ CTCs ($p = 0.0237$). The median counts of EpCAM⁺ and EpCAM⁻ CTCs were 0.00 (0.00–1.25) and 0.83 (0.00–3.32) cells/ml, respectively.

EpCAM⁺ CTCs exhibited significantly higher expression of the chemokine receptors CXCR4 and CXCR7 compared to EpCAM⁻ CTCs (25/45 vs. 6/45, $p < 0.001$). Specifically, CXCR4/7-positive cells predominated among EpCAM⁺ CTCs, whereas CXCR4/7-negative cells were more common among EpCAM⁻ CTCs. The median number of CXCR4/7-positive cells was 0.83 (0.00–2.12) cells/ml for EpCAM⁺ CTCs and 0.00 (0.00–0.00) cells/ml for EpCAM⁻ CTCs. Conversely, the median number of CXCR4/7-negative cells was 0.00 (0.00–0.00) cells/ml in EpCAM⁺ CTCs and 0.00 (0.00–0.83) cells/ml in EpCAM⁻ CTCs.

Moreover, the number of CXCR4/7-expressing CTCs was significantly higher among EpCAM⁺ CTCs compared to EpCAM⁻ CTCs ($p < 0.001$), with median values of 0.83 (0.00–2.12) cells/ml and 0.00 (0.00–0.00) cells/ml, respectively.

Analysis of the number of EpCAM⁺ and EpCAM⁻ CTCs expressing CXCR4/7, in relation to clinicopathological parameters, revealed no significant associations with clinical variables (Supplementary, Fig. 1A–D).

During the 6-year follow-up period, disease progression was observed in 3 patients: one experienced tumor recurrence, while two developed distant metastases. These patients were grouped as having tumor progression, whereas the remaining patients were classified as without progression. Comparative analysis of the number of EpCAM⁺ and EpCAM⁻ CTCs expressing CXCR4 and CXCR7 demonstrated a significant increase in the total population of EpCAM⁻ CTCs, as well as in both EpCAM⁻CXCR4/7⁺ and EpCAM⁻CXCR4/7⁻ CTC subsets in patients with progression ($p = 0.0007$, $p = 0.0184$, and $p = 0.0013$, respectively) (Fig. 1A). No significant differences were detected in the number of EpCAM⁺ CTCs, regardless of CXCR4/7 expression, between patients with and without signs of progression during the follow-up period ($p > 0.05$) (Fig. 1B).

ROC analysis of EpCAM⁻ CTC counts considering CXCR4/7 expression in patients with progression over a 6-year follow-up demonstrated the prognostic value of both EpCAM⁻ and EpCAM⁻CXCR4/7⁻ parameters (Fig. 2). An EpCAM⁻ CTC count > 2.23 cells/ml of peripheral blood predicted progression with 100.0% sensitivity and 95.1% specificity (AUC=0.96, 95% CI: 0.91–1.00; $p = 0.008$). Similarly, an EpCAM⁻CXCR4/7⁻CTC count >1.25 cells/ml predicted progression with 100.0% sensitivity and 85.7% specificity (AUC = 0.96, 95% CI: 0.89–1.00; $p = 0.009$). These results indicate that the EpCAM⁻CXCR4/7⁻ CTC population has prognostic significance, whereas the EpCAM⁻CXCR4/7⁺ cells do not (AUC = 0.80, 95% CI: 0.47–1.00; $p = 0.089$).

Analysis of stemness and EMT markers in EpCAM⁻CXCR4/7⁻ CTCs associated with tumor progression revealed the presence of stemness features based on the expression of CD44/CD24, CD133, and ALDH1A1. The number of EpCAM⁻CXCR4/7⁻CD44⁺CD24⁻CD133⁺ALDH1A1⁺ CTCs was significantly higher in patients with progression during the observation period compared to those without progression ($p = 0.003$). The median count of CTCs exhibiting the CD44⁺CD24⁻CD133⁺ALDH1A1⁺ phenotype was 0.00 (0.00–0.00) cells/ml in patients without
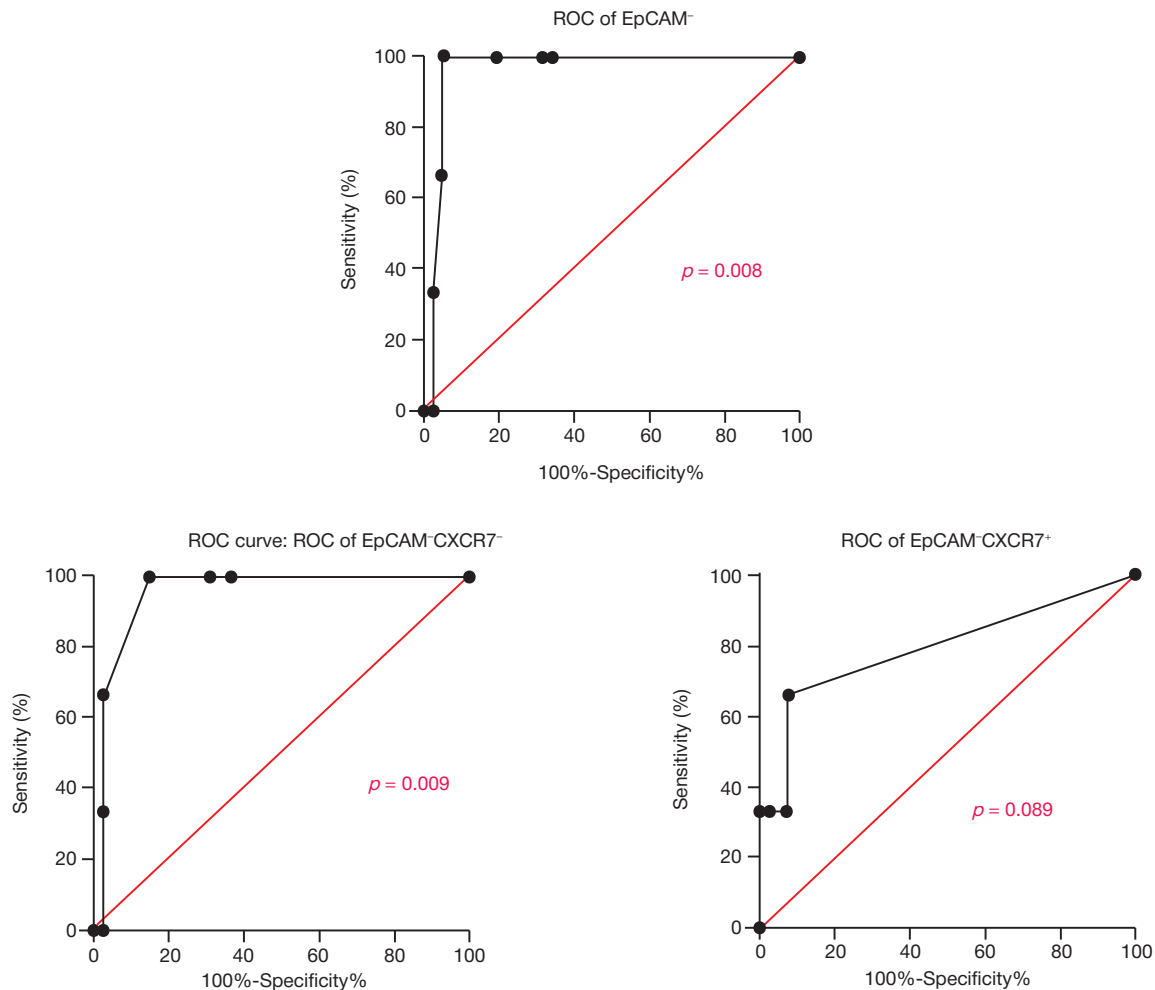
Fig. 2. ROC analysis of EpCAM$^+$ CTCs according to CXCR4/7 expression in breast cancer patients with tumor progression during the 6-year follow-up period

progression and 1.66 (0.00–1.68) cells/ml in patients with progression.

The assessment features of EMT revealed the expression of N-cadherin and Snail (p = 0.003). The median number of CTCs with the N-cadherin+Snail+Vimentin– phenotype in patients without and with the presence of signs of progression was 0.00 (0.00–0.00) cells/ml and 2.24 (0.00–4.98) cells/ml, respectively.

Association of chemokine receptor CXCR4 and CXCR7 expression with stemness traits among EpCAM$^+$ and EpCAM$^-$ CTCs

To evaluate the association between stemness features and the expression of chemokine receptors CXCR4 and CXCR7, the frequency and number of CTCs expressing stemness markers CD44/CD24, CD133, and ALDH1 were analyzed among EpCAM$^+$ and EpCAM$^-$ CTCs, considering CXCR4/7 expression. Among EpCAM$^+$ CTCs, the highest frequency was observed in cells with the CD44$^+$CD24$^-$CD133$^+$ALDH1$^+$ phenotype. Specifically, the frequency of these cells was 51.1% (23/45) when CXCR4/7 was expressed, compared to 11.1% (5/45) in its absence. Thus, the occurrence of CD44$^+$CD24$^-$CD133$^+$ALDH1$^+$ cells was significantly higher in EpCAM$^+$ CTCs expressing CXCR4/7 chemokine receptors (p < 0.0001).

The greatest number of cells were also characterized by the CD44$^+$CD24$^-$CD133$^+$ALDH$^+$ phenotype, regardless of CXCR4/7 expression (Fig. 3A). The number of cells with this stem phenotype was higher among EpCAM$^+$CXCR4/7$^+$ CTCs

(p < 0.0001). The median CD44$^+$CD24$^-$CD133$^+$ALDH$^+$ among EpCAM$^+$CXCR4/7$^+$ CTCs was 0.56 (0.00–1.67) cells/ml and 0.00 (0.00–0.00) cells/ml among EpCAM$^+$CXCR4/7$^-$ CTCs. Among EpCAM$^-$CXCR4/7$^+$ CTCs, the frequency of occurrence and the number of cells with stemness variants did not differ (p > 0.05) (Fig. 3B).

The frequency of EpCAM$^-$CXCR4/7$^-$ CTCs was extremely low. The most common phenotype among these cells was CD44$^-$CD24$^-$CD133$^-$ALDH$^-$, observed in 12 out of 45 cases, while the frequency of all other phenotypes did not exceed 2 out of 45 (p = 0.02). Additionally, the number of CD44$^-$CD24$^-$CD133$^-$ALDH$^-$ cells was significantly higher compared to other phenotypes, with significance levels indicated in the figure above. The median number of cells with this phenotype was 0.00 (0.00–0.70) cells/ml (Fig. 4B).

Evaluation of chemokine receptors CXCR4 and CXCR7 expression in EpCAM$^+$ CTCs with distinct stemness phenotypes revealed that a significantly higher proportion of CD44$^+$CD24$^-$CD133$^+$ALDH1$^+$ cells expressed CXCR4/7 (p < 0.0001) (Fig. 4A). In contrast, the number of CTCs with or without CXCR4/7 expression did not differ significantly among cells exhibiting other stemness phenotypes (p > 0.05).

Analysis of chemokine receptors CXCR4 and CXCR7 expression in EpCAM$^-$ CTCs, considering stemness characteristics, showed no significant differences in CXCR4/7 expression among stem-like CTCs (Fig. 4B). However, a significantly greater number of cells lacking stemness features, characterized by the CD44$^-$CD24$^-$CD133$^-$ALDH$^-$ phenotype, were negative for CXCR4/7 expression (p = 0.0011).
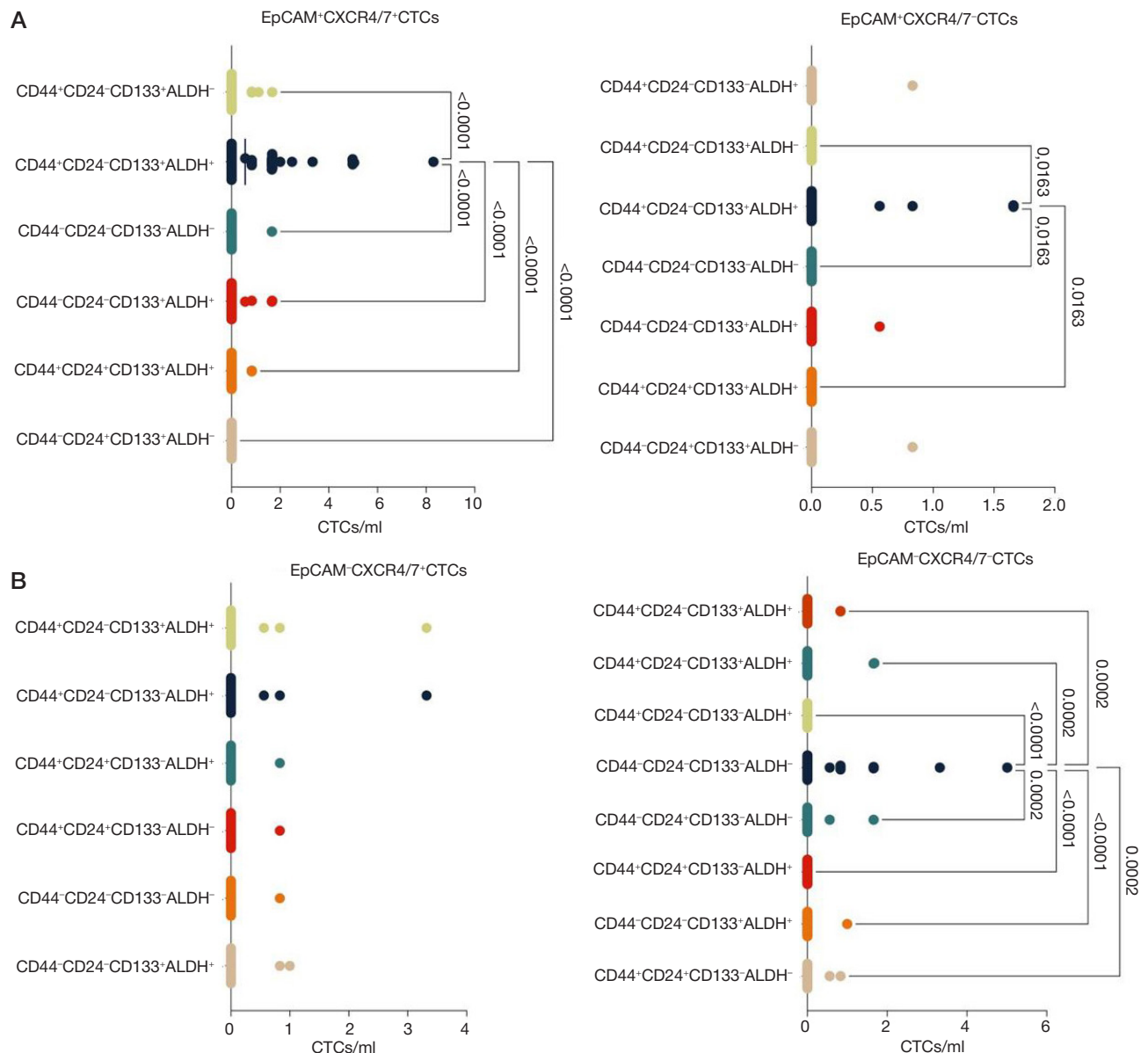
Fig. 3. Number of EpCAM$^+$ (**A**) and EpCAM$^-$ (**B**) CTCs considering CXCR4/7 expression with stem features in breast cancer patients

Association of chemokine receptor CXCR4 and CXCR7 expression with EMT features among EpCAM$^+$ and EpCAM$^-$ CTCs

In EpCAM$^+$ and EpCAM$^-$ CTCs, the expression of the early EMT marker Snail, which represses epithelial markers and promotes mesenchymal marker expression, was evaluated alongside late EMT markers N-cadherin and vimentin. Among EpCAM$^+$CXCR4/7$^+$ CTCs, the highest frequency was observed in cells expressing the late EMT markers N-cadherin and vimentin, accounting for 35.6% (16/45). Furthermore, the number of N-cadherin$^+$Snail$^-$vimentin$^+$ cells was significantly higher compared to both N-cadherin$^+$Snail$^-$vimentin$^-$ and N-cadherin$^-$Snail$^-$vimentin$^-$ CTCs ($p = 0.0003$ and $p = 0.0009$, respectively) (Fig. 5A). In contrast, among EpCAM$^+$ CTCs lacking CXCR4/7 expression, the distribution of cells with different EMT phenotypes did not differ significantly ($p > 0.05$). In this case, N-cadherin$^+$Snail$^-$Vimentin$^+$ cells were practically not found among CXCR4/7$^-$ CTCs; their number was higher among CXCR4/7$^+$ CTCs ($p < 0.0001$) (Fig. 5B).

Analysis of early and late EMT marker expression in EpCAM$^-$ CTCs, considering CXCR4/7 expression, revealed no significant differences ($p > 0.05$) (Fig. 6A). Notably, among EpCAM$^-$ CTCs

expressing CXCR4/7, no cells were found to co-express all three EMT markers analyzed.

Among EpCAM$^-$ CTCs lacking N-cadherin and expressing Snail, regardless of vimentin status, a significantly greater number of cells were negative for CXCR4/7 expression ($p = 0.0061$ and $p = 0.0189$, respectively) (Fig. 6B).

Expression of CXCR4 and CXCR7 genes in *EPCAM$^+$* and *EPCAM$^-$* CTCs

In 20 BC patient samples analyzed, a total of 239 CTCs were identified. *EPCAM$^-$* CTCs were defined as cells lacking *PTPRC* (CD45) and EPCAM gene expression but exhibiting positive expression of cytokeratin genes (*KRT7*, *KRT8*, or *KRT18*). Conversely, *EPCAM$^+$* CTCs were characterized as cells without *PTPRC* (CD45) expression and with *EPCAM* gene expression levels greater than zero, regardless of cytokeratins expression. Consequently, the *EPCAM$^+$* and *EPCAM$^-$* CTC groups comprised 11 and 228 cells, respectively. The frequency of SDF-1 chemokine receptor gene expression — *CXCR4* or *CXCR7* — did not differ significantly between *EPCAM$^+$* and *EPCAM$^-$* CTCs, being 54.5% (6/11) and 53.5% (122/228), respectively.
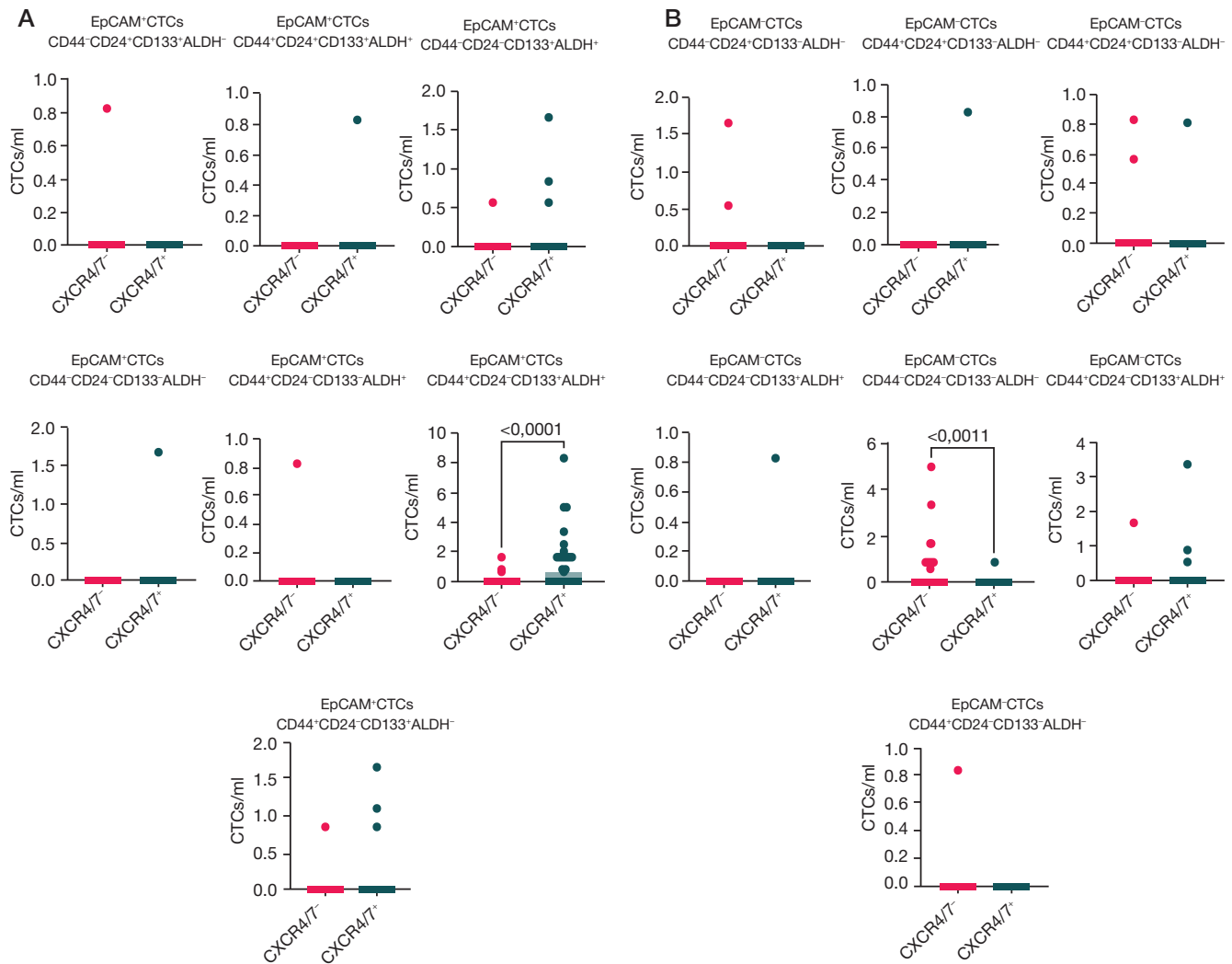
Fig. 4. Comparison of the number of EpCAM⁻ (**A**) and EpCAM⁺ (**B**) CTCs with stemness properties according to CXCR4/7 expression in breast cancer

Differential gene expression analysis among *EPCAM⁺* cells with and without *CXCR4/7* expression revealed no significant differences ($p > 0.05$). In contrast, comparison within *EPCAM⁻* CTCs showed significant differences. The most overexpressed genes in *EPCAM⁻ CXCR4/7⁺* CTCs included *POSTN* ($p = 6.63 \times 10^{-13}$), *FN1* ($p = 5.24 \times 10^{-14}$), *COL3A1* ($p = 1.08 \times 10^{-11}$), *VIM* ($p = 2.33 \times 10^{-19}$), *S100A6* ($p = 4.60 \times 10^{-9}$), and *CD74* ($p = 1.11 \times 10^{-7}$). According to the KEGG 2021 Human database, the ribosome metabolism pathway showed the highest enrichment of overexpressed genes ($p = 2.365 \times 10^{-143}$), while the MSigDB Hallmark 2020 database identified the Myc Targets V1 pathway as the most upregulated in *EPCAM⁻ CXCR4/7⁺* cells ($p = 7.59 \times 10^{-11}$) (Supplementary, Fig. 2A).

In the *EPCAM⁻CXCR4/7⁻* CTC population, increased expression of numerous genes was observed, with the most significantly overexpressed being *PF4* ($p = 1.33 \times 10^{-24}$), *PPBP* ($p = 5.40 \times 10^{-23}$), and *TUBB1* ($p = 1.62 \times 10^{-20}$). According to the KEGG 2021 Human database, the largest group of upregulated genes was associated with the ferroptosis pathway ($p = 3.315 \times 10^{-7}$), while the MSigDB Hallmark 2020 database highlighted the androgen receptor signaling pathway as significantly enriched ($p = 8.0 \times 10^{-5}$) (Supplementary, Fig. 2B).

Comparison of the transcriptional profiles between *EPCAM⁺* and *EPCAM⁻* CTCs expressing *CXCR4/7* revealed no significant differences ($p > 0.05$).

During the observation period starting in 2020, disease progression was observed in 4 out of 20 patients. Among these,

3 patients developed metastases to distant organs, while one patient exhibited metastasis to regional lymph nodes. The distribution of CTC subpopulations in individual patient samples is summarized in the Supplementary, Table 1.

No significant differences were observed in the frequency and number of *EPCAM⁺* and *EPCAM⁻* CTCs expressing *CXCR4/7* chemokine receptor genes between patients with different treatment outcomes over the 6-year follow-up period ($p > 0.05$). Additionally, differential gene expression analysis was performed on *EpCAM⁻ CXCR4/7⁻* cells from patients with and without disease progression during the follow-up. In the group of patients with tumor progression, three genes showed significantly increased expression: *HBB* ($p = 1.34 \times 10^{-5}$), *IGLC2* ($p = 7.49 \times 10^{-6}$), and *IGHM* ($p = 1.05 \times 10^{-5}$). In contrast, among patients without progression, only one gene, *MALAT1*, was significantly overexpressed ($p = 1.52 \times 10^{-2}$).

Spatial transcriptomic analysis of *EPCAM⁺* and *EPCAM⁻* tumor cells in relation to *CXCR4* and *CXCR7* gene expression

Manual annotation of spots in five BC samples was conducted to identify those containing tumor cells. Spots exclusively featuring stromal cells or spots where the number of stromal cells surpassed that of tumor cells were excluded from the analysis. Subsequently, employing the Gene Filter tool, all spots were categorized into two groups based on *EPCAM* gene
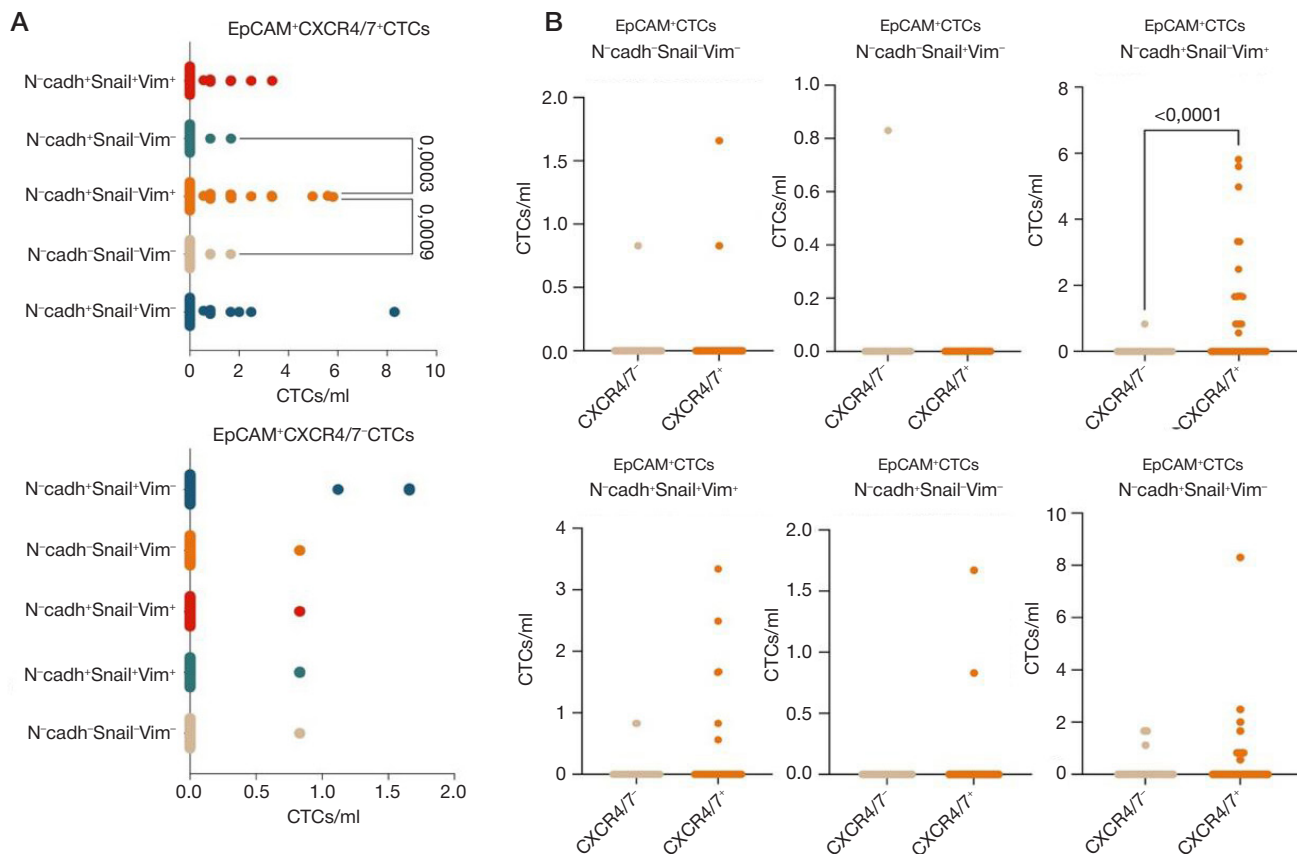
Fig. 5. Number of EpCAM+ CTCs considering CXCR4/7 expression with EMT features in breast cancer patients

expression levels. Spots with *EPCAM* expression ≤ 2 units were designated as *EPCAM−*, spots with *EPCAM* expression ≥ 3 units were classified as *EPCAM+*. Within each group of *EPCAM+* and *EPCAM−* spots, the expression of the chemokine receptor genes *CXCR4* and *CXCR7* was assessed, leading to the identification of clusters comprising spots negative for both genes (*CXCR4/7−*) and clusters containing spots expressing at least one of the two genes (*CXCR4/7+*). The transcriptional profiles of *EPCAM+* spots expressing *CXCR4* and/or *CXCR7* were compared to those lacking *CXCR4/7* expression. *EPCAM+CXCR4/7+* spots exhibited a substantial number of differentially expressed genes, with the top 100 listed in Supplementary, Table 3. Most upregulated genes were associated with estrogen signaling pathways ($p = 0.0039$) and cell-cell or cell-matrix adhesion processes ($p = 0.0039$) (Supplementary, Fig. 3A). While, *EPCAM+CXCR4/7−* spots showed activation of endocytosis ($p = 0.0359$) and an early response to estrogen ($p = 1.435 \times 10^{-7}$) (Supplementary, Fig. 3B). Comparison of *EPCAM−* spots based on *CXCR4/7* chemokine receptor gene expression revealed significant activation of EMT ($p = 1.6223 \times 10^{-58}$) and protein digestion and absorption pathway ($p = 5.723 \times 10^{-13}$) in *EPCAM−CXCR4/7+* tumor cells (Supplementary, Fig. 4A). (Supplementary, Table 5). Additionally, in *EPCAM−CXCR4/7+* tumor cells, the largest number of overexpressed genes were associated with the early response to estrogen signature ($p = 1.442 \times 10^{-13}$) (Supplementary, Fig. 4B).

We also compared the transcriptional profiles of *EPCAM+* and *EPCAM−* tumor cells expressing *CXCR4* and/or *CXCR7*. The analysis revealed activation of the G2-M checkpoint signature in *EPCAM+ CXCR4/7+* tumor cells ($p = 1.086 \times 10^{-8}$) (Supplementary, Fig. 5A).

In contrast, *EPCAM−CXCR4/7+* tumor cells were characterized by activation of the protein digestion and absorption signature ($p = 4.231 \times 10^{-16}$) and EMT ($p = 3.952 \times 10^{-30}$) (Supplementary, Fig. 5B).

DISCUSSION

The findings of this study reveal a profound heterogeneity in CTCs of BC patients, with distinct phenotypic and transcriptional signatures that correlate with clinical outcomes. By integrating flow cytometry and transcriptomic data, we provided a comprehensive view of CTC subpopulations according chemokine receptors CXCR4 and CXCR7 expression and their functional implications.

There are relatively few studies in the literature focusing on chemokine receptor expression in CTCs. Notably, Mego et al. (2016) isolated CTCs from peripheral blood using CD45-negative selection, followed by RT-PCR to evaluate the expression of target genes [19]. The authors identified CTCs expressing epithelial markers (*KRT19*) as well as mesenchymal markers (*TWIST1*, *SNAIL1*, *SLUG*, and *ZEB1*) and further characterized the gene expression of various chemokine receptors within these populations. Their findings demonstrated that epithelial *KRT19+* CTCs exhibited higher expression levels of the *CXCR4* receptor, and its ligand SDF-1 compared to mesenchymal CTCs. In our study, flow cytometry analysis demonstrated a clear dichotomy in CXCR4/7 expression between different subpopulations of epithelial CTCs. While CXCR4/7-positive cells were predominant among EpCAM+ CTCs, the EpCAM− population was enriched for CXCR4/7-negative cells. This segregation suggests that these markers define distinct subpopulations of epithelial CTCs with different biological behaviors.

It is important to note that tumor cells expressing CXCR4 are frequently associated with cancer stem cells in the literature [20]. Indeed, CXCR4+ tumor cells exhibit key stem cell characteristics, including a high proliferation rate, resistance to conventional therapies, and enhanced metastatic potential [21]. However, our study did not find a correlation between
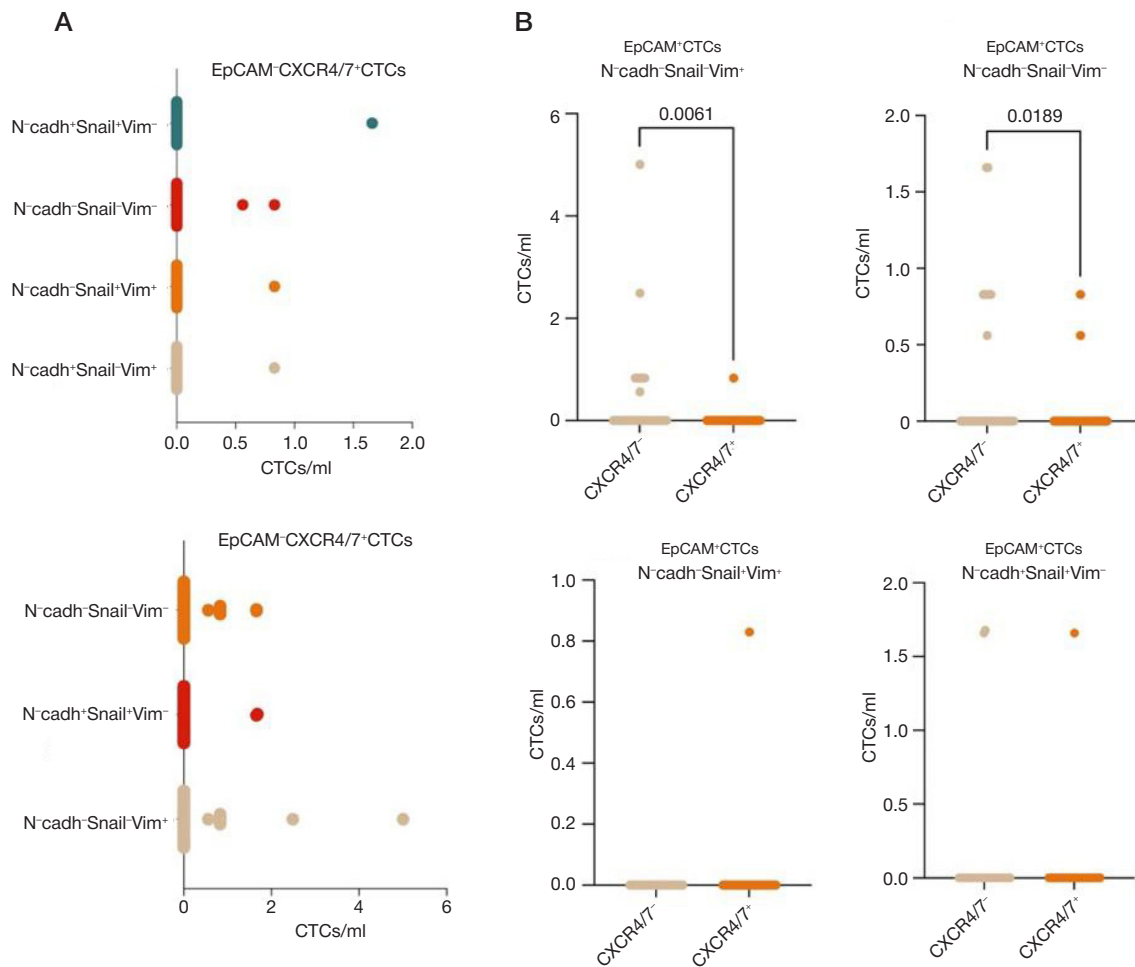
Fig. 6. Number of EpCAM⁺ CTCs considering CXCR4/7 expression with EMT features in breast cancer patients

stemness features and the expression of CXCR4/7 chemokine receptors in either the EpCAM⁺ or EpCAM⁻ CTC populations. However, stem cell features characteristic of the EpCAM⁺ and EpCAM⁻ CTC subpopulations were identified. EpCAM⁺ CTCs were predominantly characterized by a stem cell phenotype defined as CD44⁺CD24⁻CD133⁺ALDH⁺, whereas EpCAM⁻ CTCs largely consisted of cells lacking stemness markers (CD44⁻CD24⁻CD133⁻ALDH⁻).

Analysis of the expression of early (Snail) and late (N-cadherin and vimentin) EMT markers revealed a correlation between N-cadherin and vimentin expression and CXCR4/7 chemokine receptor presence in EpCAM⁻ CTCs. The N-cadherin⁺Snail⁻vimentin⁺ phenotype was predominantly observed in the EpCAM⁻ CXCR4/7⁺ CTC subpopulation. This finding aligns with existing evidence indicating that activation of the CXCL12/CXCR4 signaling axis can induce EMT in breast cancer cells via stimulation of the Wnt/⁻-catenin and mTOR signaling pathways [22].

Our findings highlight the necessity to move beyond EpCAM-based CTC detection. Thus, the analysis of CXCR4/7 protein expression in CTCs enabled the identification of an association between EpCAM⁻CXCR4/7⁻ CTCs and tumor progression, suggesting the potential prognostic value of this subpopulation. While EpCAM-expressing cells have traditionally been regarded as the primary drivers of metastasis, accumulating evidence highlights the importance of EpCAM⁻ subpopulations [5]. Loss of EpCAM expression is frequently linked to EMT, with the hybrid EMT phenotype — characterized by the concurrent expression of epithelial and mesenchymal markers — being considered the most aggressive and metastatic [23]. Consistent with this, the EpCAM⁻CXCR4/7⁻ CTC population identified in our study

exhibited expression of EMT markers such as N-cadherin and Snail. Surprisingly, the EpCAM⁻CXCR4/7⁻ subset — rather than the expected CXCR4/7⁺ population — emerged as the most prognostically significant, indicating that metastatic potential may not solely depend on chemokine receptor-driven dissemination pathways.

The lack of prognostic value in EpCAM⁻CXCR4/7⁺ CTCs (despite statistical significance in Mann–Whitney tests) raises important questions. These cells may represent a transient or dormant state, where CXCR4/7 signaling facilitates survival in circulation but does not directly drive metastatic outgrowth. In contrast, the EpCAM⁻CXCR4/7⁻ subset may harbor more aggressive, immune-evasive clones that bypass conventional detection methods yet drive progression. Clarification of the stem and EMT traits in the detected population of CTCs showed that progression was associated with cells characterized by stemness based on the expression of CD44/CD24, CD133 and ALDH1A1 markers, as well as those with the EMT phenotype — N-cadherin⁺Snail⁺vimentin⁻. The obtained results are consistent with the data of the world literature indicating a high metastatic potential of tumor cells with signs of stemness and EMT [24].

Transcriptomic analysis of CTCs uncovered striking differences between subpopulations. EPCAM⁻CXCR4/7⁺ CTCs exhibited marked overexpression of genes associated with extracellular matrix (ECM) remodeling and stromal activation (POSTN, FN1, COL3A1, VIM, S100A6, CD74), suggesting a role in premetastatic niche formation. In contrast, EPCAM⁻CXCR4/7⁻ CTCs displayed upregulation of PF4, PPBP, and TUBB1, genes linked to platelet and microtubule dynamics,

potentially indicating alternative mechanisms of dissemination. Notably, in patients with tumor progression, HBB, IGLC2, and IGHM were significantly overexpressed, possibly reflecting immune evasion or clonal selection, whereas MALAT1 was the sole gene elevated in non-progressors, consistent with its known role in tumor suppression. The HBB gene, which encodes beta-globin — a key component of hemoglobin — exhibits a complex and context-dependent role in breast cancer. Although HBB expression is traditionally associated with erythrocytes, it has also been detected in breast cancer cells, where its function appears to be dualistic. Some studies report that elevated HBB expression correlates with increased tumor aggressiveness, enhanced metastatic potential, and poorer patient prognosis [25]. Conversely, other research suggests that HBB may exert tumor-suppressive effects under specific conditions in certain cancer types [26]. Two others differentially expressed genes identified are associated with immunoglobulin synthesis. The IGLC2 gene encodes the constant region of the immunoglobulin lambda light chain (Immunoglobulin lambda constant 2), which is involved in antigen binding. To date, the only available study linking IGLC2 expression to breast cancer indicates its role as a predictor of a favorable clinical outcome in the triple-negative breast cancer subtype [27]. In contrast, no data currently exist regarding the association of IGHM expression, which encodes the constant region of the immunoglobulin M heavy chain, with tumor growth or progression. In the group of patients without signs of progression, only one gene, MALAT1, was found to be overexpressed. MALAT1 is a long non-coding RNA associated with metastasis in lung adenocarcinoma. Its function is linked to the regulation of cell motility and invasive potential [28]. Notably, MALAT1 has also been reported to suppress breast cancer metastasis [29]. Specifically, the study demonstrated that MALAT1 can bind to the pro-metastatic transcription factor TEAD, inactivating it and thereby inhibiting tumor cell migration and invasion. Furthermore, the authors observed that MALAT1 expression is frequently reduced in more aggressive and metastatic breast cancer forms, supporting its role as a metastasis suppressor.

Spatial transcriptomics (Visium 10X) of primary breast tumors further corroborated these findings, revealing that EPCAM⁻CXCR4/7⁺ regions were enriched for ECM-related genes (COL1A1, COL3A1, FN1, POSTN, SPARC, BGN), indicative of a fibrotic, immune-modulated microenvironment. Conversely, EPCAM⁻CXCR4/7⁻ regions overexpressed a wide range of genes, among which STC2, TFF3, NPNT, and CD24 were the most functionally significant in our opinion, suggesting features associated with an aggressive phenotype. Despite the controversial association with the prognosis, recent data show, that secreted STC2 functions as a ligand in an autocrine/paracrine manner to promote cell survival by alleviating oxidative stress [30]. Cancer tissue expression of Trefoil factor 3 (TFF3) are identified as prognostic indicators of dormant ER⁺ BC with TFF3 functioning as an epigenetically regulated driver of dormancy-associated behaviors [31]. The experiment demonstrated that knockdown of NPNT reduced the adhesion of cancer cells to osteoblasts, confirming its role in bone metastasis in breast cancer [32]. Several genes in the EpCAM⁻CXCR4/7⁻ regions (CDH1, CRABP2, THSD4, SERPINA1, SERPINA3, HSPB1, KRT8, CD9, NUPR1, AZGP1) have been associated with suppression of migration and invasion. However, considering the existence of intravasation mechanisms that do not require invasion, and in light of the data above, this may indicate a functional metastatic phenotype characteristic of this population. Probably, this phenotype is capable of withstanding the effects of an aggressive environment while remaining dormant, with the ability to adapt within a premetastatic niche.

These results challenge the conventional view of CTC biology, as the EpCAM⁻ subset emerged as a key predictor of progression. Notably, EpCAM⁻CXCR4/7⁺ CTCs exhibit mesenchymal and extracellular matrix remodeling characteristics, whereas the CXCR4/7⁻ subset may harbor more aggressive clones. Further functional studies are required to elucidate the mechanistic roles of these CTC subsets in metastasis and therapy resistance.

Despite the significant findings, this study has several important limitations. The primary objective was to characterize CTCs based on chemokine receptor CXCR4 and CXCR7 expression and their role in receptor-driven dissemination pathways. However, the prospective study design limited patient recruitment, resulting in only three cases with disease progression. While this small sample size precludes definitive conclusions, the results provide a valuable foundation for future research aimed at identifying pathogenetically relevant CTC subpopulations. Notably, our reanalysis of single-cell transcriptomic data confirmed the functional profiles of CTCs associated with progression. Further support comes from spatial transcriptomics data, which revealed CXCR4/7-associated heterogeneity within the primary tumor, consistent with our CTC findings and underscoring their biological relevance. It is important to emphasize that this observational study identifies associations between CTC phenotypes and clinical outcomes that require mechanistic validation *in vitro* and *in vivo*. Additionally, spatial transcriptomics, while informative about the tumor microenvironment, has limited resolution (~55 μm), potentially resulting in signal averaging across different cell types.

However, these limitations do not diminish the significance of the findings and instead highlight the need for further studies with larger cohorts, employing single-cell analysis and functional experiments to validate the identified patterns. The current data provide a solid foundation for expanded investigations into the role of CXCR4/7-expressing CTCs in disease progression.

CONCLUSIONS

Taking into account the limitations of this study, particularly the small sample size including patients with tumor progression in follow-up period, several conclusions can be drawn. Tumor progression, characterized by tumor cell dissemination to distant organs, may not be directly associated with the presence of CXCR4 and CXCR7 receptors on CTCs. At the same time, the EpCAM⁻ CTC population appears to be pathogenetically significant for tumor progression. The number of EpCAM⁻ CTCs, irrespective of CXCR4 and CXCR7 expression, was higher in patients exhibiting progression during the follow-up period. This finding underscores the need to shift the focus of CTC research from EpCAM⁺ CTCs — which have shown limited prognostic value in early breast cancer over more than two decades — to the EpCAM⁻ subpopulation. Transcriptomic analysis of EPCAM⁻CXCR4/7⁻ CTCs revealed distinct gene expression profiles; however, their precise role in breast cancer progression remains inadequately understood. Considering both quantitative and qualitative alterations in these cells, it is plausible that patients with poor prognosis are characterized not only by an increased number of EPCAM⁻CXCR4/7⁻CTCs but also by changes in their functional properties.

Data availability

The datasets analyzed in this study are available in the Gene Expression Omnibus (GEO) Database under the accession number GSE242311 and in the BioProject under the accession number PRJNA776403.

References

1. Thomas-Bonafos T, Pierga JY, Bidard FC, Cabel L, Kiavue N, et al. Circulating tumor cells in breast cancer: clinical validity and utility. NPJ Breast Cancer. 2024; 10: 103.
2. Ju S, Chen C, Zhang J, Xu L, Zhang X, Li Z, Chen Y, Zhou J, Ji F, Wang L, et al. Detection of circulating tumor cells: opportunities and challenges. Biomark Res. 2022; 10: 58.
3. Eslami-S Z, Cortés-Hernández LE, Alix-Panabières C. Epithelial Cell Adhesion Molecule: An Anchor to Isolate Clinically Relevant Circulating Tumor Cells. Cells. 2020; 9: 1836.
4. Riethdorf S, Fritsche H, Müller V, Rau T, Schindlbeck C, Rack B, Janni W, Coith C, Beck K, Jänicke F, et al. Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: A validation study of the CellSearch system. Clin Cancer Res. 2007; 13: 920–28.
5. Ye F, Zhong X, Qiu Y, Yang L, Wei B, Zhang Z, Bu H. The presence of EpCAM⁻/CD49f⁺ cells in breast cancer is associated with a poor clinical outcome. J Breast Cancer. 2015; 18: 242–8.
6. Nieto MA, Huang RY, Jackson RA, Thiery JP. Cell. 2016; 166: 21–45.
7. Aiello NM, Kang Y. Context-dependent EMT programs in cancer metastasis. J Exp Med. 2019; 216: 1016–26.
8. Rueda A, Serna N, Mangues R, Vázquez E, Villaverde A. Targeting the chemokine receptor CXCR4 for cancer therapies. Biomark Res. 2025; 13: 68.
9. Chatterjee S, Behnam Azad B, Nimmagadda S. The intricate role of CXCR4 in cancer. Adv Cancer Res. 2014; 124: 31–82.
10. Boldajipour B, Mahabaleshwar H, Kardash E, Reichman-Fried M, Blaser H, Minina S, et al. Control of chemokine-guided cell migration by ligand sequestration. Cell. 2008; 132: 463–73.
11. Yang Y, Li J, Lei W. CXCL12-CXCR4/CXCR7 Axis in Cancer: from Mechanisms to Clinical Applications. Int J Biol Sci. 2023; 19: 3341–59.
12. Wu C, Zhao H, Chen H, Yao Q. CXCR4 in breast cancer: oncogenic role and therapeutic targeting. Drug Des Devel Ther. 2015; 9: 4953–64.
13. Pospelova RA. Leukocyte concentration in clinical practice: diagnostic value. M.: Meditsina, 1973.
14. Menyailo M, Zainullina V, Khozyainova A, Tashireva L, Zolotareva S, Gerashchenko T, et al. Heterogeneity of Circulating Epithelial Cells in Breast Cancer at Single-Cell Resolution: Identifying Tumor and Hybrid Cells. Adv Biol (Weinh). 2023; 7: e2200206.
15. Satija Lab. Seurat: Tools for Single Cell Genomics. GitHub. 2023. Available from: https://github.com/satijalab/seurat.
16. McGinnis CS. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data. GitHub. 2023. Available from: https://github.com/chris-mcginnis-ucsf/DoubletFinder.
17. Tashireva L, Grigoryeva E, Alifanov V, et al. Spatial Heterogeneity of Integrins and Their Ligands in Primary Breast Tumors. Discov Med. 2023; 35: 910–20.
18. Hafemeister C, Satija R. Normalization and variance stabilization

19. Mego M, Cholujova D, Minarik G, Sedlackova T, Gronesova P, Karaba M, et al. CXCR4-SDF-1 interaction potentially mediates trafficking of circulating tumor cells in primary breast cancer. BMC Cancer. 2016; 16: 127T.
20. Shi Q, Yang W, Ouyang Y, Zhou S, Xu L, Liu J, et al. CXCR4 promotes tumor stemness maintenance and CDK4/6 inhibitors resistance in ER-positive breast cancer. Breast Cancer Res. 2025; 27: 15.
21. Nengroo MA, Maheshwari S, Singh A, Verma A, Arya RK, Chaturvedi P, et al. CXCR4 intracellular protein promotes drug resistance and tumorigenic potential by inversely regulating the expression of Death Receptor 5. Cell Death Dis. 2021; 12: 464.
22. Garg P, Jallepalli VR, Verma S. Unravelling the CXCL12/CXCR4 Axis in breast cancer: Insights into metastasis, microenvironment interactions, and therapeutic opportunities. Hum Gene. 2024; 40: 201272.
23. Jolly MK, Mani SA, Levine H. Hybrid epithelial/mesenchymal phenotype(s): The 'fittest' for metastasis? Biochim Biophys Acta Rev Cancer. 2018; 1870: 151–7.
24. Fabregat I, Malfettone A, Soukupova J. New Insights into the Crossroads between EMT and Stemness in the Context of Cancer. J Clin Med. 2016; 5: 37.
25. Ponzetti M, Capulli M, Angelucci A, Delle Monache S, Sica AR, Festuccia C, et al. Non-conventional role of haemoglobin beta in breast malignancy. Br J Cancer. 2017; 117: 994–1006.
26. Zhang X, Yang H, Zhang L, Li Y, Ma H. A Pan-Cancer Analysis of the Hemoglobin Subunit Beta (HBB) in Human Tumors. Front Genet. 2022; 13: 880647.
27. Chang YT, Tsai WC, Lin WZ, Chen YS, Hou MF, Wu YC, et al. A Novel IGLC2 Gene Linked With Prognosis of Triple-Negative Breast Cancer. Front Oncol. 2022; 11: 759952.
28. Tsyganov MM, Ibragimova MK. MALAT1 Long Non-coding RNA and Its Role in Breast Carcinogenesis. Acta Naturae. 2023; 15: 32–41.
29. Kim J, Piao HL, Kim BJ, Yao F, Han Z, Wang Y, et al. Long noncoding RNA MALAT1 suppresses breast cancer metastasis. Nat Genet. 2018; 50: 1705–15.
30. Qie S, Xiong H, Liu Y, et al. Stanniocalcin 2 governs cancer cell adaptation to nutrient insufficiency through alleviation of oxidative stress. Cell Death Dis. 2024; 15: 567. https://doi.org/10.1038/s41419-024-06961-7.
31. Chen S, Zhang X, Basappa B, et al. TFF3 facilitates dormancy of anti-estrogen treated ER+ mammary carcinoma. Commun Med. 2025; 5: 45. https://doi.org/10.1038/s43856-024-00710-9.
32. Wang D, Zhao C, Gao L, Wang Y, Gao X, Tang L, et al. NPNT promotes early-stage bone metastases in breast cancer by regulation of the osteogenic niche. J Bone Oncol. 2018; 13: 91–96. DOI: 10.1016/j.jbo.2018.09.006.

of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019; 20: 296.

Литература

1. Thomas-Bonafos T, Pierga JY, Bidard FC, Cabel L, Kiavue N, et al. Circulating tumor cells in breast cancer: clinical validity and utility. NPJ Breast Cancer. 2024; 10: 103.
2. Ju S, Chen C, Zhang J, Xu L, Zhang X, Li Z, Chen Y, Zhou J, Ji F, Wang L, et al. Detection of circulating tumor cells: opportunities and challenges. Biomark Res. 2022; 10: 58.
3. Eslami-S Z, Cortés-Hernández LE, Alix-Panabières C. Epithelial Cell Adhesion Molecule: An Anchor to Isolate Clinically Relevant Circulating Tumor Cells. Cells. 2020; 9: 1836.
4. Riethdorf S, Fritsche H, Müller V, Rau T, Schindlbeck C, Rack B, Janni W, Coith C, Beck K, Jänicke F, et al. Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: A validation study of the CellSearch system. Clin Cancer Res. 2007; 13: 920–28.
5. Ye F, Zhong X, Qiu Y, Yang L, Wei B, Zhang Z, Bu H. The presence of EpCAM⁻/CD49f⁺ cells in breast cancer is associated with a poor clinical outcome. J Breast Cancer. 2015; 18: 242–8.
6. Nieto MA, Huang RY, Jackson RA, Thiery JP. Cell. 2016; 166: 21–45.
7. Aiello NM, Kang Y. Context-dependent EMT programs in cancer metastasis. J Exp Med. 2019; 216: 1016–26.
8. Rueda A, Serna N, Mangues R, Vázquez E, Villaverde A. Targeting the chemokine receptor CXCR4 for cancer therapies. Biomark Res. 2025; 13: 68.
9. Chatterjee S, Behnam Azad B, Nimmagadda S. The intricate role of CXCR4 in cancer. Adv Cancer Res. 2014; 124: 31–82.
10. Boldajipour B, Mahabaleshwar H, Kardash E, Reichman-Fried M, Blaser H, Minina S, et al. Control of chemokine-guided cell migration by ligand sequestration. Cell. 2008; 132: 463–73.
11. Yang Y, Li J, Lei W. CXCL12-CXCR4/CXCR7 Axis in Cancer: from Mechanisms to Clinical Applications. Int J Biol Sci. 2023; 19: 3341–59.
12. Wu C, Zhao H, Chen H, Yao Q. CXCR4 in breast cancer: oncogenic role and therapeutic targeting. Drug Des Devel Ther. 2015; 9: 4953–64.

13. Pospelova RA. Leukocyte concentration in clinical practice: diagnostic value. M.: Meditsina, 1973.
14. Menyailo M, Zainullina V, Khozyainova A, Tashireva L, Zolotareva S, Gerashchenko T, et al. Heterogeneity of Circulating Epithelial Cells in Breast Cancer at Single-Cell Resolution: Identifying Tumor and Hybrid Cells. Adv Biol (Weinh). 2023; 7: e2200206.
15. Satija Lab. Seurat: Tools for Single Cell Genomics. GitHub. 2023. Available from: https://github.com/satijalab/seurat.
16. McGinnis CS. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data. GitHub. 2023. Available from: https://github.com/chris-mcginnis-ucsf/DoubletFinder.
17. Tashireva L, Grigoryeva E, Alifanov V, et al. Spatial Heterogeneity of Integrins and Their Ligands in Primary Breast Tumors. Discov Med. 2023; 35: 910–20.
18. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019; 20: 296.
19. Mego M, Cholujova D, Minarik G, Sedlackova T, Gronesova P, Karaba M, et al. CXCR4-SDF-1 interaction potentially mediates trafficking of circulating tumor cells in primary breast cancer. BMC Cancer. 2016; 16: 127T.
20. Shi Q, Yang W, Ouyang Y, Zhou S, Xu L, Liu J, et al. CXCR4 promotes tumor stemness maintenance and CDK4/6 inhibitors resistance in ER-positive breast cancer. Breast Cancer Res. 2025; 27: 15.
21. Nengroo MA, Maheshwari S, Singh A, Verma A, Arya RK, Chaturvedi P, et al. CXCR4 intracellular protein promotes drug resistance and tumorigenic potential by inversely regulating the expression of Death Receptor 5. Cell Death Dis. 2021; 12: 464.
22. Garg P, Jallepalli VR, Verma S. Unravelling the CXCL12/CXCR4 Axis in breast cancer: Insights into metastasis, microenvironment interactions, and therapeutic opportunities. Hum Gene. 2024; 40: 201272.
23. Jolly MK, Mani SA, Levine H. Hybrid epithelial/mesenchymal phenotype(s): The 'fittest' for metastasis? Biochim Biophys Acta Rev Cancer. 2018; 1870: 151–7.
24. Fabregat I, Malfettone A, Soukupova J. New Insights into the Crossroads between EMT and Stemness in the Context of Cancer. J Clin Med. 2016; 5: 37.
25. Ponzetti M, Capulli M, Angelucci A, Delle Monache S, Sica AR, Festuccia C, et al. Non-conventional role of haemoglobin beta in breast malignancy. Br J Cancer. 2017; 117: 994–1006.
26. Zhang X, Yang H, Zhang L, Li Y, Ma H. A Pan-Cancer Analysis of the Hemoglobin Subunit Beta (HBB) in Human Tumors. Front Genet. 2022; 13: 880647.
27. Chang YT, Tsai WC, Lin WZ, Chen YS, Hou MF, Wu YC, et al. A Novel IGLC2 Gene Linked With Prognosis of Triple-Negative Breast Cancer. Front Oncol. 2022; 11: 759952.
28. Tsyganov MM, Ibragimova MK. MALAT1 Long Non-coding RNA and Its Role in Breast Carcinogenesis. Acta Naturae. 2023; 15: 32–41.
29. Kim J, Piao HL, Kim BJ, Yao F, Han Z, Wang Y, et al. Long noncoding RNA MALAT1 suppresses breast cancer metastasis. Nat Genet. 2018; 50: 1705–15.
30. Qie S, Xiong H, Liu Y, et al. Stanniocalcin 2 governs cancer cell adaptation to nutrient insufficiency through alleviation of oxidative stress. Cell Death Dis. 2024; 15: 567. https://doi.org/10.1038/s41419-024-06961-7.
31. Chen S, Zhang X, Basappa B, et al. TFF3 facilitates dormancy of anti-estrogen treated ER+ mammary carcinoma. Commun Med. 2025; 5: 45. https://doi.org/10.1038/s43856-024-00710-9.
32. Wang D, Zhao C, Gao L, Wang Y, Gao X, Tang L, et al. NPNT promotes early-stage bone metastases in breast cancer by regulation of the osteogenic niche. J Bone Oncol. 2018; 13: 91–96. DOI: 10.1016/j.jbo.2018.09.006.