Research Square

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

Context-dependent dual role of GRHL1 in triple negative and luminal B breast cancer

Emma Spreemann
University Hospital Aachen, RWTH Aachen University
Celasty Das
University Hospital Aachen, RWTH Aachen University
Thomas Hansen
University Hospital Aachen, RWTH Aachen University
Elisabeth Schygulla
University of Freiburg
Hannah Heidenreich
University Hospital Aachen, RWTH Aachen University
Lena Hahnfeld
University Hospital Aachen, RWTH Aachen University
Michael Rose
University of Ulm
Peter Bronsert
University of Freiburg
Elmar Stickeler
University Hospital Aachen, RWTH Aachen University
Jochen Maurer
jmaurer@ukaachen.de

University Hospital Aachen, RWTH Aachen University

Research Article

Keywords: GRHL1, triple-negative breast cancer, breast cancer stem cells, human epidermal growth factor receptor 2, stemness, migration

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11425569/v1

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

Additional Declarations: No competing interests reported.

Context-dependent dual role of GRHL1 in triple negative and luminal B breast cancer

Emma Spreemann [1], Celasty Das [1], Thomas Hansen [1], Elisabeth Schygulla [2,3,4], Hannah Heidenreich [1], Lena Hahnfeld [1], Michael Rose [5,6], Peter Bronsert [2], Elmar Stickeler [1] and Jochen Maurer [1,*]

[1] Department of Obstetrics and Gynecology, University Hospital Aachen, RWTH Aachen University, 52074 Aachen, Germany

[2] Institute for Surgical Pathology, Medical Center, University of Freiburg, 79106 Freiburg, Germany

[3] Faculty of Medicine, University of Freiburg, 79106 Freiburg, Germany

[4] Comprehensive Cancer Center Freiburg, Medical Center, University of Freiburg, 79106 Freiburg, Germany

[5] Institute of Pathology, University Hospital, University of Ulm, 89081 Ulm, Germany

[6] Institute of Pathology, University Hospital Aachen, RWTH Aachen University, 52074 Aachen, Germany

* Corresponding author: Jochen Maurer, jmaurer@ukaachen.de; Tel.: +49-0241-80-37051

Abstract

Background: Breast cancer remains the leading cause of female cancer mortality worldwide. Triple-negative breast cancer is associated with a poor prognosis and limited therapeutic options due to the absence of well-defined molecular targets. Grainyhead-like 1 (GRHL1), an evolutionarily conserved transcription factor regulating epithelial differentiation, displays context-dependent tumor-suppressive or oncogenic roles.

Methods: Given the pivotal role of breast cancer stem cells in chemotherapy resistance, metastasis, and recurrence, we investigated the functional role of GRHL1 in primary patient derived breast cancer stem cells and publicly available breast cancer cell lines.

Results: Breast cancer subtype analysis reveals that GRHL1 mRNA expression is elevated in the Basal-like subtype. Notably, elevated GRHL1 correlates with prolonged overall survival specifically within triple-negative breast cancer, whereas lower expression predicts favorable outcomes in other subtypes. Depletion of GRHL1 in triple-negative breast cancer cell lines enhanced self-renewal while reducing migratory capacity and epithelial–mesenchymal transition marker expression. Conversely, GRHL1 overexpression suppressed self-renewal and increased migration and epithelial–mesenchymal transition markers. In a luminal B breast cancer stem cell line, GRHL1 exerted opposing effects, with overexpression enhancing stemness features and depletion leading to a reduction thereof.

Conclusions: These findings emphasize the context-dependent dual role of GRHL1, highlighting its subtype-specific influence on breast cancer progression and plasticity.

Keywords

GRHL1, triple-negative breast cancer, breast cancer stem cells, human epidermal growth factor receptor 2, stemness, migration

Introduction

Breast cancer is the most diagnosed cancer in women as well as one of the leading causes for cancer deaths worldwide (1). Especially triple-negative breast cancer (TNBC), which is defined by the absence of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor (HER2), the key structures for targeted therapies, has a poor prognosis (2). TNBC is a heterogeneous disease, which is more likely to resist therapies, reoccur after treatment, and metastasize. These challenges in treating TNBC are partly due to a higher fraction of breast cancer stem cells (BCSCs) in the tumor (3, 4). BCSCs exist in a highly flexible state of epithelial-mesenchymal plasticity (EMP), being able to both self-renew and activate the differentiation into more epithelial or mesenchymal phenotypes (5). The process of epithelial-mesenchymal transition (EMT) is linked to promote migration, invasion, and stem-like properties. Conversely, the reverse process mesenchymal-epithelial transition (MET) is associated with tumor-initiation ability, a requirement for metastatic colonization (6, 7).

The Grainy head (Grh) transcription factor family was first described in Drosophila melanogaster as a crucial regulator of development and integrity of epidermal barriers. Named Grainyhead-like (GRHL) transcription factor family in vertebrates, it is evolutionary highly conserved over different species and contains three members GRHL1, GRHL2 and

GRHL3 (8, 9). Recent studies of the GRHL family show tumor suppressive as well as oncogenic effects dependent on the GRHL family member, the cancer entity or even its subtype (10, 11).

Recent studies using an ectopic expression of GRHL2 in breast cancer (12) showed reduced mammosphere formation, suppressed the cancer stem cell phenotype (CD44+/CD24-) and showed reduced tumor initiation in xenograft models, apparently through suppression of EMT (12, 13). Overexpression (OE) of GRHL3 in breast cancer increases migration and invasion through repression of the E-cadherin promoter activity (14).

The function of GRHL1 as well as its target genes in breast cancer remains unknown (15). Dependent on the tumor entity, GRHL1 can act as tumor suppressor or oncogene (10). In non-small cell lung cancer, higher GRHL1 expression was associated with poor patient survival and promoted cell growth possibly through cell cycle progression (16). Similarly, endometrial cancer showed increased GRHL1 expression in comparison to normal tissue, high GRHL1 expression correlates with worse overall as well as progression free survival and patients with higher levels of GRHL1 were more resistant to immunotherapy (17). In contrast, high expression of GRHL1 in neuroblastoma correlates with favorable patient survival and reexpression of GRHL1 in MYCN-amplified neuroblastoma cells reduced proliferation, colony formation and xenograft growth (18). In squamous cell carcinoma of the skin GRHL1 acts as tumor suppressor. GRHL1 knockout mice showed increased tumor-initiation with earlier onset through deregulated terminal differentiation of keratinocytes with mild chronic inflammation (19). Since no previous studies explored the

role of GRHL1 in breast cancer, we aimed to investigate its expression and function with a focus on TNBC and beyond that, the luminal B breast cancer subtype.

Results

1. GRHL1 mRNA expression affects patient overall survival dependent on breast cancer subtype

To investigate the general role of GRHL1 on breast cancer we first analyzed the overall survival of breast cancer patients from the Cancer Genome Atlas (TCGA) Firehose data, which were classified into high and low GRHL1 expression cohorts, according to the median expression. We discovered that lower GRHL1 mRNA expression correlates with favorable overall survival in breast cancer in general (Figure 1 A). Due to the heterogeneity of breast cancer, we further divided the patient cohorts by their receptor status. ER+ and PR+ patient breast cancer tumors showed a low GRHL1 mRNA expression whereas TNBC tumors exhibited a high level of GRHL1 mRNA expression (Table 1). In each group, ER- and ER+ as well as PR- and PR+, low mRNA expression of GRHL1 was correlated to a better overall survival for breast cancer patients (Additional File 1 A, B). In HER2 negative breast cancer, the GRHL1 expression does not correlate with the overall survival of patients (Figure 1 B), while low expression of GRHL1 in HER2+ breast cancer patients correlates with a significantly better overall survival (Figure 1 C). These results suggest a relationship between GRHL1 expression and the HER2 status. Similarly, the non-TNBC group showed better overall survival with low GRHL1 mRNA expression (Figure 1 D). In contrast to that, high GRHL1 expression in TNBC was associated with a trend towards more favorable overall survival (Figure 1 E). This group in general has 3 times more

patients with higher GRHL1 expression than low GRHL1 expression (Table 1, Figure 1 E). Furthermore, we investigated the potential effect of GRHL1 in different stages of breast cancer and divided the patients into stage I and II versus stage III and IV breast cancer (AJCC). In higher stage breast cancer, lower GRHL1 expression significantly predicts better overall survival (Additional File 1 C). To analyze the differences between adjacent normal tissue and tumor tissue, we investigated the GRHL1 mRNA expression through matched samples. The tumor tissue expressed significantly less GRHL1 in comparison to normal tissue (Figure 1 F). Due to the differences we found between receptor-defined subgroups (ER, PR, HER2) in survival, we next investigated the GRHL1 mRNA expression in each molecular subtype. The basal-like subtype exceeds all other subtypes, showing a significantly higher mRNA expression than all other subtypes (Figure 1 G).

Table 1. Clinicopathological parameters in relation to GRHL1 mRNA-expression of breast cancer patients from TCGA Firehose data.

		GRHL1 mRNA expression[a]			
Parameter	**n**	**low**	**high**	**p-value[b]**	**Spearman Correlation[c]**
Total	1091	545	546		
Age at Diagnosis					
< 58	518	262	256	0.673	0.013
≥ 58	572	282	290		
Tumor size					
T1	810	407	403	0.781	0.008
≥ T1	278	137	141		
Lymph Node status					
pN0	556	277	279	0.779	-0.009
pN1-3	515	261	254		

Tumor stage					
Stage I-II	800	391	409	0.152	-0.044
Stage III-IV	267	144	123		
Estrogen receptor status					
Negative	235	61	174	**<0.001**	-0.262
Positive	806	462	344		
Progesterone receptor status					
Negative	340	108	232	**<0.001**	-0.259
Positive	698	414	284		
HER2 status					
Negative	559	275	284	0.614	0.019
Positive	164	77	87		
Triple-negative status					
Non-TNBC	929	497	432	**<0.001**	0.186
TNBC	114	27	87		

Significant P-values are marked in **boldface**. Abbreviations: HER2, human epidermal growth factor 2; TNBC, triple-negative breast cancer.

[a] Dichotomized at the median.

[b] Pearson´s chi-Square test (two-sided) was used due to the large sample size.

[c] Spearman´s rank correlation coefficient was used to assess correlations.

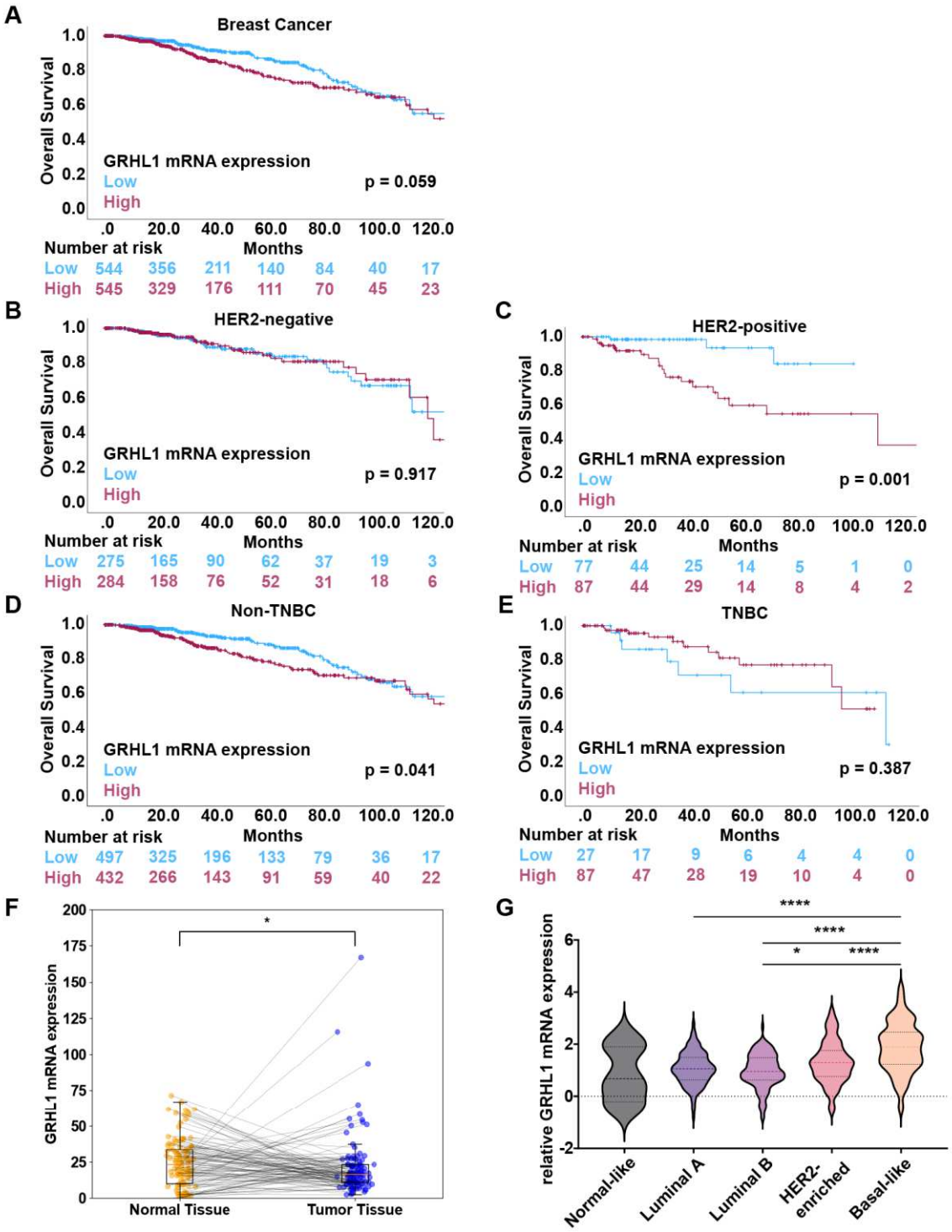


Figure 1. GRHL1 mRNA expression affects patient overall survival dependent on breast cancer subtype. A, B Kaplan-Meier survival analysis of GRHL1 mRNA-sequencing TCGA Firehose data. GRHL1 mRNA-sequencing data were separated by median in two groups of low GRHL1 mRNA expression (blue curve, n = 544) and high

GRHL1 mRNA expression (red curve, n = 545). Patients were divided in (**B**) HER2- and (**C**) HER2+ and overall survival between GRHL1 low and high mRNA expression groups is displayed (pooled p-value = 0.103). **D, E** Patients were split into (**D**) non-TNBC and (**E**) TNBC, and overall survival data is shown (pooled p-value = 0.071). Statistical significance was analyzed using log-rank test. **F** GRHL1 mRNA expression in tumor tissue matched with normal tissue (n = 120). Statistical significance was determined using Wilcoxon test. **G** Relative GRHL1 mRNA expression in molecular subtypes (Normal-like n = 8, Luminal A n = 231, Luminal B n = 127, HER2-enriched n = 58 Basal-like n = 96). Data was analyzed using Ordinary one-way ANOVA with Holm- Šídák correction. Mean values ± SD, n = 3. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

2. High GRHL1 expression in tumor micro array correlates with better overall survival in TNBC patients

Nuclear GRHL1 expression was assessed immunohistochemically in 16 patients with triple-negative breast cancer. They were stratified into low and high GRHL1 expression groups according to the median H-score of nuclear staining, with eight patients assigned to each group. The Kaplan Meier analysis showed a non-significant ($p = 0.056$) trend towards longer overall survival in patients with high nuclear GRHL1 expression compared with those with low nuclear GRHL1 expression (Figure 2 A). In univariable Cox regression analysis, high nuclear GRHL1 expression was also associated with a lower hazard ratio for death compared with low GRHL1 expression (Table 2). However, this association represents only a statistical trend, but did not reach significance (HR = 0.23, 95% CI: 0.05–1.17, $p = 0.077$). Also, no statistically significant associations were observed

in overall survival for pathological tumor stage, lymph node stage, tumor grading or age in univariable analysis (Table 2). Further a multivariable Cox regression analysis, was performed for pathological tumor stage, lymph node stage, tumor grading and age. In this model, high nuclear GRHL1 expression was not significantly associated with overall survival (HR = 0.28, 95% CI: 0.02–5.06, p = 0.4).

Table 2. Univariable und multivariable Cox regression analysis of overall survival in patients with triple-negative breast cancer

	Univariable Cox				**Multivariable Cox**			
Characteristic	**N**	**HR**	**95% CI**	**p-value**	**N**	**HR**	**95% CI**	**p-value**
	16				12			
low		—	—			—	—	
high		0.23	0.05, 1.17	0.077		0.28	0.02, 5.06	0.4
	16				12			
pT1 - 2		—	—			—	—	
pT3 - 4		2.70	0.62, 11.7	0.2		22.4	0.82, 611	0.065
	12				12			
pN0		—	—			—	—	
pN1		1.43	0.15, 13.8	0.8		3.08	0.19, 49.0	0.4
pN2 - 3		2.46	0.49, 12.3	0.3		0.78	0.04, 15.3	0.9
	15				12			
G1		—	—					

	Univariable Cox				Multivariable Cox			
Characteristic	**N**	**HR**	**95% CI**	**p-value**	**N**	**HR**	**95% CI**	**p-value**
G2		1.09	0.26, 4.56	> 0.9		—	—	
G3						0.20	0.02, 1.89	0.2
Age	16				12			
< 70		—	—			—	—	
> 70		0.88	0.22, 3.54	0.9		0.56	0.06, 5.30	0.6

Abbreviations: CI, Confidence Interval; HR, Hazard Ratio

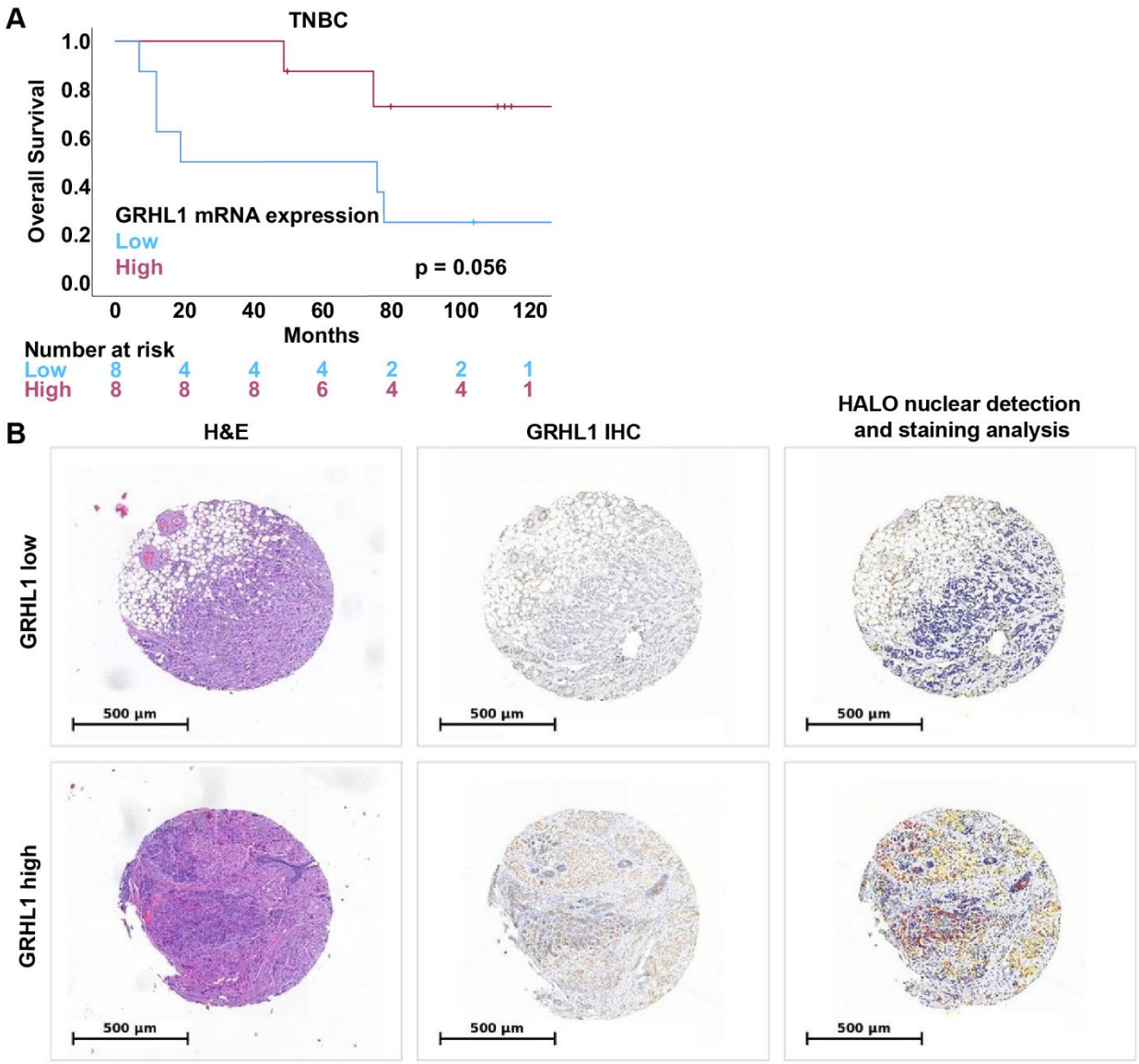


Figure 2. High GRHL1 expression in tumor micro array correlates with better overall survival in TNBC patients. A Kaplan-Meier analysis of overall survival according to nuclear GRHL1 expression (GRHL1 low n = 8; GRHL1 high n = 8) in triple-negative breast cancer patients. Statistical significance was analyzed using log-rank test. **B** Representative images of two cores with haematoxylin and eosin (H&E) staining as well as corresponding GRHL1 immunohistochemistry and HALO based nuclear detection and staining analysis (blue = no staining, yellow = weak staining, orange = moderate staining,

red = strong staining). The detected nuclei are grouped into tumor, stroma, healthy glands, fat cells and lymphocytes according to a classifier trained in HALO AI. Scale bars represent 500 µm.

3. Knockdown and overexpression of GRHL1 does not affect proliferation and morphology of TNBC cell lines

Since GRHL1 expression seems to be particularly linked to the TNBC subtype we further focused our efforts on this breast cancer subtype. To examine general GRHL1 expression of our previously established cancer cell lines (20-22) we performed RT-qPCR and western blot for the GRHL1 expression in our BCSCs, BCSC1-5 being TNBC and BCSC6 being luminal B (ER 3% positive, PR negative, HER2neu 1+ (low-positive)) breast cancer. BCSC2 and BCSC6 showed the highest GRHL1 expression while BCSC1, BCSC4 and BCSC5 express a comparably lower amount of GRHL1. BCSC3 differed between its GRHL1 expression and showed discrepancies between mRNA and protein expression of GRHL1 (Figure 3 A, B). In order to get a better understanding of these results, we also investigated the GRHL1 expression of the commercial TNBC cell lines MDA-MB-468 and MDA-MB-231 as well as the immortalized “normal” breast cell line MCF10A. MDA-MB-468 showed the highest expression of those cell lines on both mRNA and protein level, while MCF10A showed higher GRHL1 expression on mRNA level and MDA-MB-231 expressed higher protein levels (Additional File 2 A, B). To investigate the biological function of GRHL1 in TNBC cell lines we transduced MDA-MB-468, BCSC2 and BCSC4 with lentiviral constructs encoding a GRHL1 CRISPR Cas9 mediated knockdown (KD). MDA-MB-468, MDA-MB-231, MCF10A, BCSC1 and BCSC5 were transduced in the

same manner with a GRHL1 overexpression (OE). The cell lines were chosen for each OE or KD based on their previously measured GRHL1 expression. GRHL1 KD and OE cell lines were validated on mRNA and protein level (Figure 3 C, D, E and Additional File 2 C, D, E). Furthermore, we analyzed the proliferation of the KD and OE cell lines. KD and OE did not affect the proliferation of neither MDA-MB-468 and MCF10A (Additional File 3) nor BCSCs (Figure 4). In addition, GRHL1 KD and OE did not change the cell morphology of the BCSCs (Additional File 4).

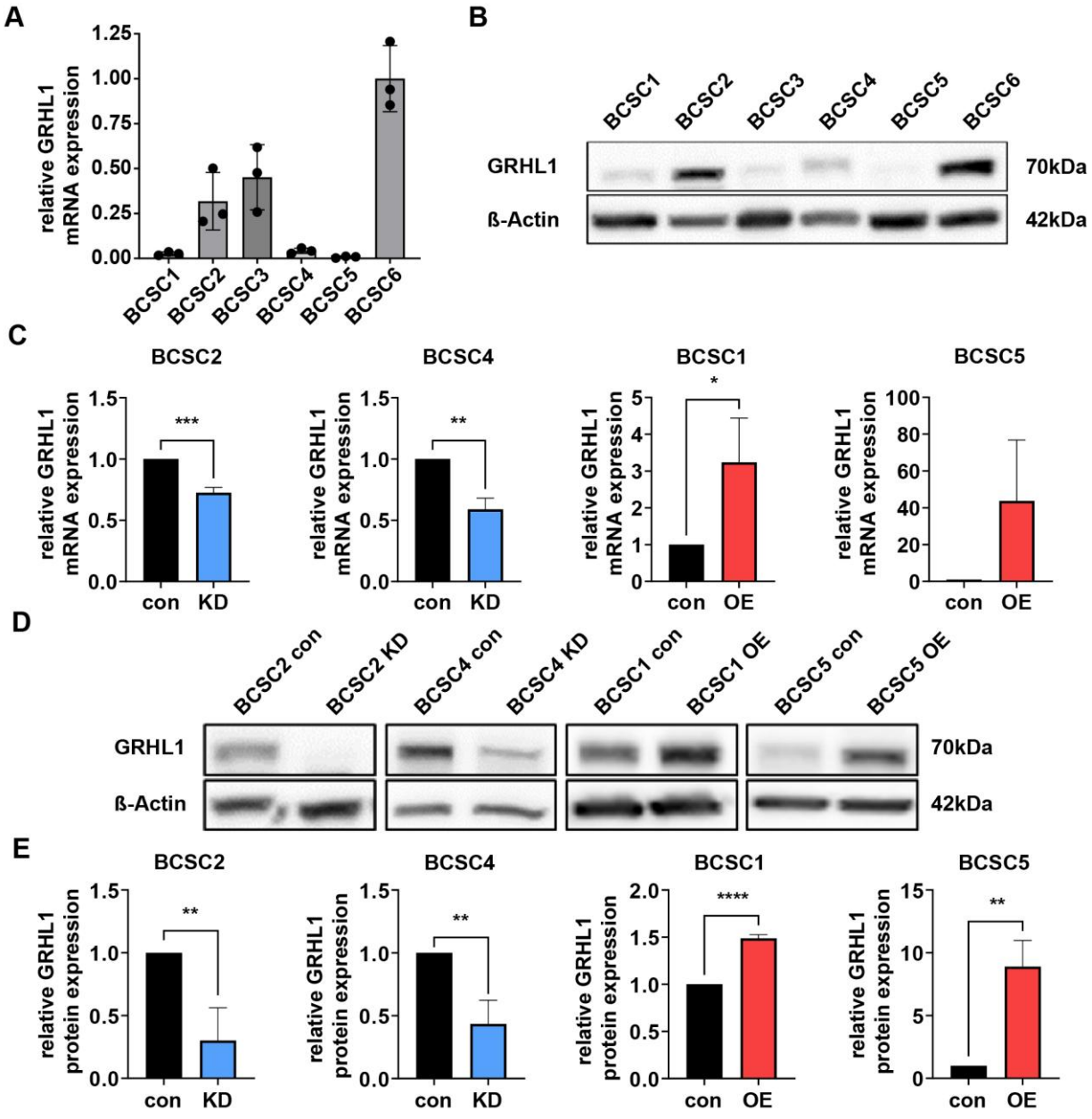


Figure 3. GRHL1 expression in BCSC and validation of GRHL1 knockdown and overexpression. A GRHL1 mRNA expression in different BCSCs lines. **B** Representative western blot of GRHL1 protein expression in different BCSCs. **C** Knockdown (KD) and overexpression (OE) of GRHL1 in TNBC cell lines was validated via reverse transcription quantitative polymerase chain reaction (RT-qPCR). **D, E** Verification of KD and OE of GRHL1 on protein level in TNBC cell lines. **D** Representative images of western blot. **E**

Quantification of western blot analysis illustrated as relative GRHL1 protein expression of KD and OE compared to the control (con: crispr scr, stuffer) in TNBC cell lines. Mean values ± SD, n = 3. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ by unpaired t-test.

4. GRHL1 regulation affects self-renewal in TNBC

Since we did not observe any effects on proliferation, we further investigated stemness and self-renewal in the manipulated cell lines, since they have been shown to be affected by GRHL1 in recent studies in other tumor entities (18, 23). We performed a sphere formation assay in the previously established KD and OE BC cell lines. GRHL1 KD in MDA-MB-468 cells increased the sphere formation significantly by 40% while OE decreased the sphere formation by 56% compared to the control. Likewise, GRHL1 OE in MDA-MB-231 cells reduced sphere formation significantly by 38% and OE of MCF10A diminished the sphere formation to 67% of the control. (Figure 5 A, Additional File 5). In BCSCs the effect of GRHL1 was similar to the effect on commercial cell lines. KD of BCSC2 increased the sphere formation significantly by 31%, while KD of BCSC4 showed a minor increase in its sphere formation capacity. OE in BCSC1 and BCSC5 significantly reduced the sphere formation to 71% in BCSC1 and 89% in BCSC5 (Figure 5 B, C).

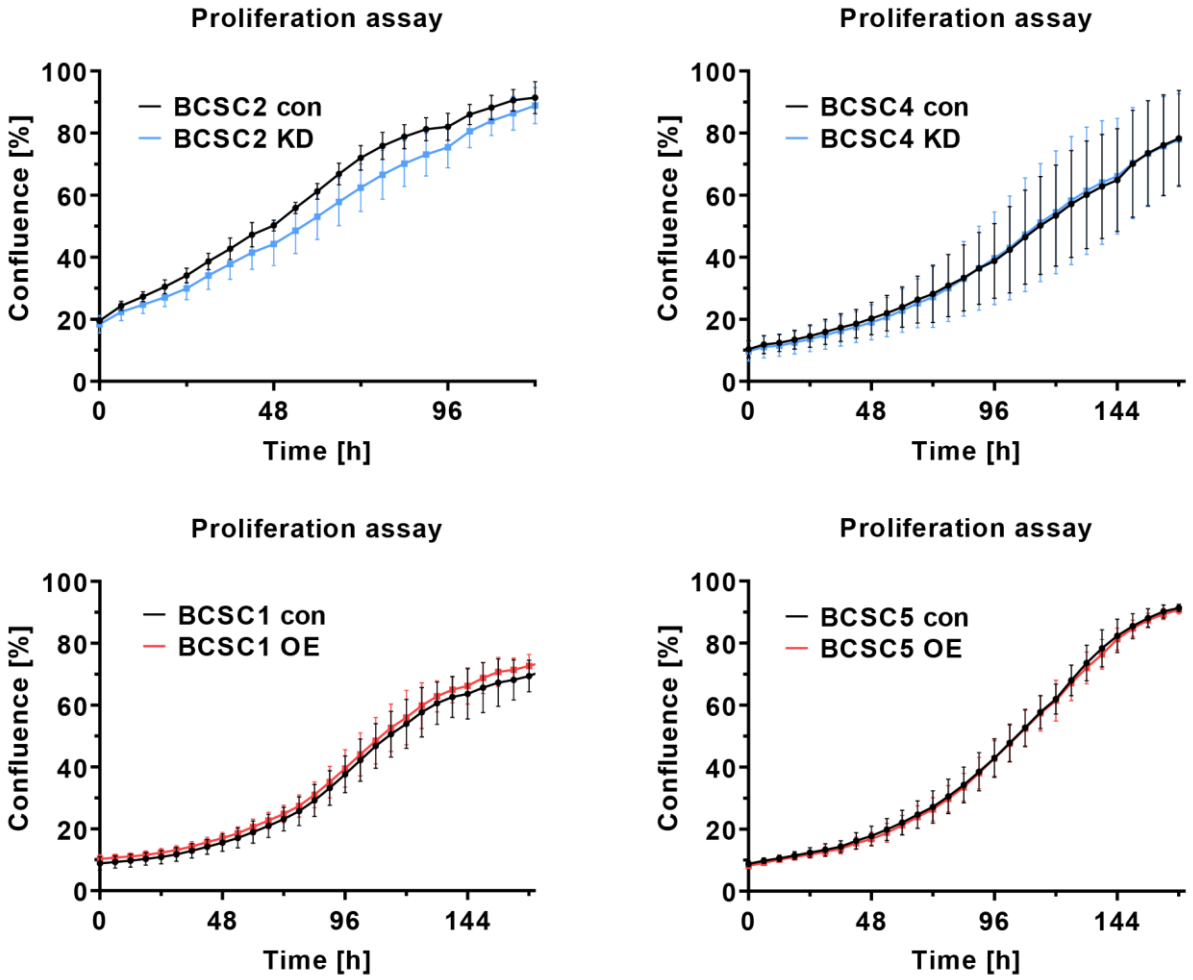


Figure 4. Manipulation of GRHL1 does not affect proliferation of TNBC cell lines. Proliferation assay comparing GRHL1 knockdown (KD) and overexpression (OE) BCSCs with control cells (con: crispr scr, stuffer). Data represents mean values ± SD, n = 3.

5. GRHL1 alteration influenced migratory behavior in TNBC cell lines

Since the GRHL transcription factor (TF) family has also been reported to affect tumor progression by enhancing or suppressing migration (18, 24, 25), we examined the effect of GRHL1 on migratory behavior, using scratch wound assays. GRHL1 KD of

BCSC2 reduced the relative wound density and thereby the ability of the cells to migrate. In addition, GRHL1 OE of BCSC5 enhanced the relative wound density. GRHL1 OE in BCSC5 increased the cells capacity to migrate (Figure 6 A, B). This effect of the change in migratory behavior was not observed in scratch wound assays using commercial cell lines. GRHL1 KD or OE of MDA-MB-468 as well as MCF10A GRHL1 OE did not alter the migratory behavior in a scratch wound assay. (Additional File 6).

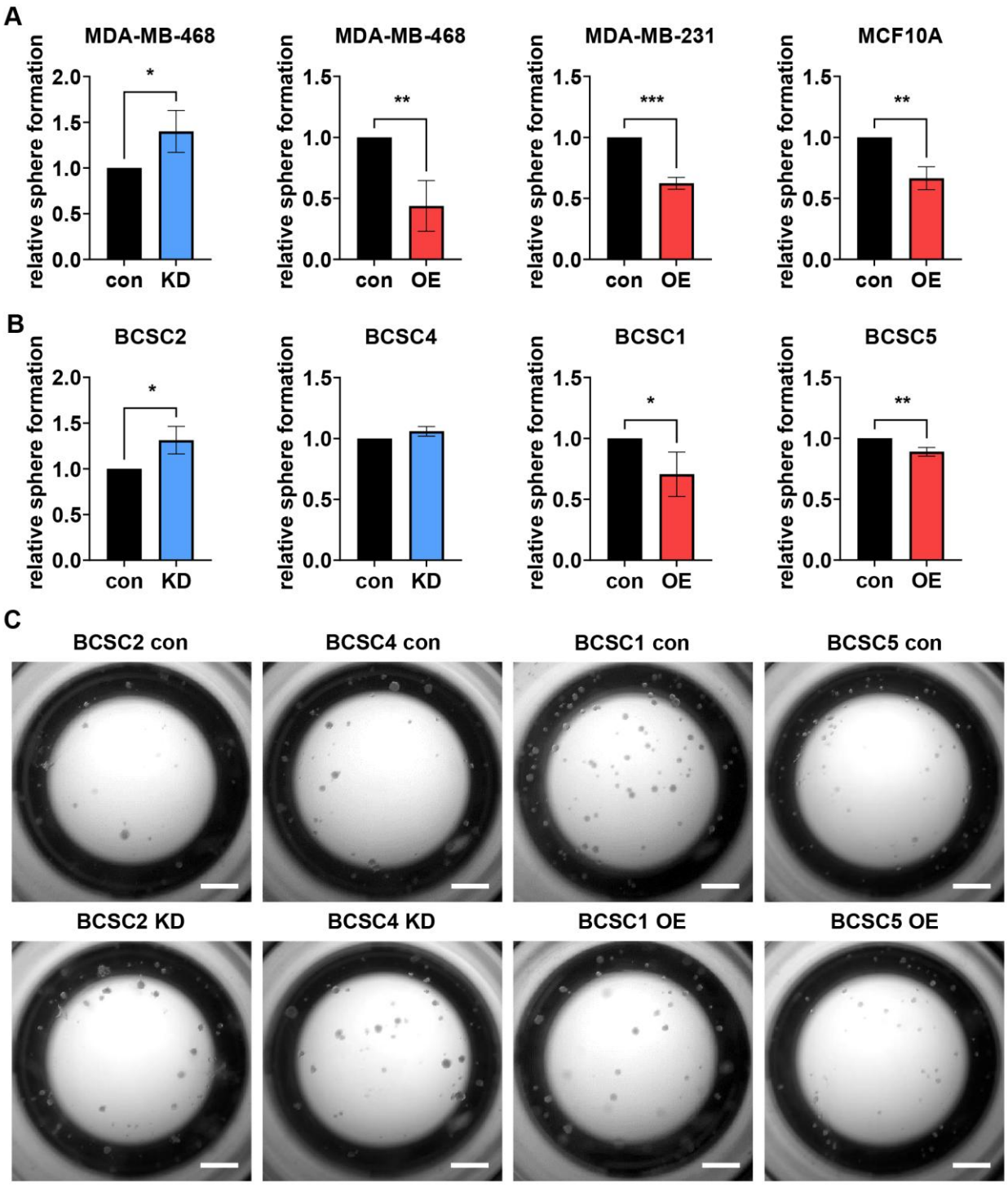


Figure 5. GRHL1 regulation affects self-renewal in TNBC. A, B Analysis of sphere formation assays of GRHL1 knockdown (KD) and GRHL1 overexpression (OE) cell lines were compared to the control (con: crispr scr, stuffer). **C** Representative whole well images

of spheres. Scale bars represent 1 mm. Data represent mean values ± SD, n = 3. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by unpaired t-test.

6. GRHL1 manipulation affects the EMT signature genes of triple-negative BCSCs

Since KD and OE of GRHL1 influenced migration and tumorigenicity, we analyzed EMT markers in BCSC using RT-qPCR. We focused on E-cadherin, N-cadherin, P-cadherin, Vimentin, Snail, Slug, KLF4 and CD44. GRHL1 KD in BCSC2 reduced the mesenchymal markers N-cadherin and Snail significantly, while Slug and CD44 were also diminished. Conversely, GRHL1 KD in BCSC2 showed a modest reduction of the epithelial marker E-cadherin as well as a significant decrease in KLF4 and P-cadherin (Figure 7 A). GRHL1 KD of BCSC4 reduced E-cadherin and the mesenchymal marker Snail significantly and KLF4. (Figure 7 B). GRHL1 KD in BCSC2 and BCSC4 had no effect on Vimentin (Figure 7 A, B). In contrast, GRHL1 OE in BCSC1 and BCSC5 showed an increase in the majority of EMT markers. GRHL1 OE of BCSC1 and BCSC5 increased epithelial markers E-cadherin, KLF4 and mesenchymal markers N-cadherin, Snail, Slug and CD44 (Figure 7 C, D). The epithelial marker P-cadherin was increased by GRHL1 OE in BCSC1 and varies in BCSC5 GRHL1 OE. Vimentin showed a modest reduction in GRHL1 OE BCSCs (Figure 7 C, D).

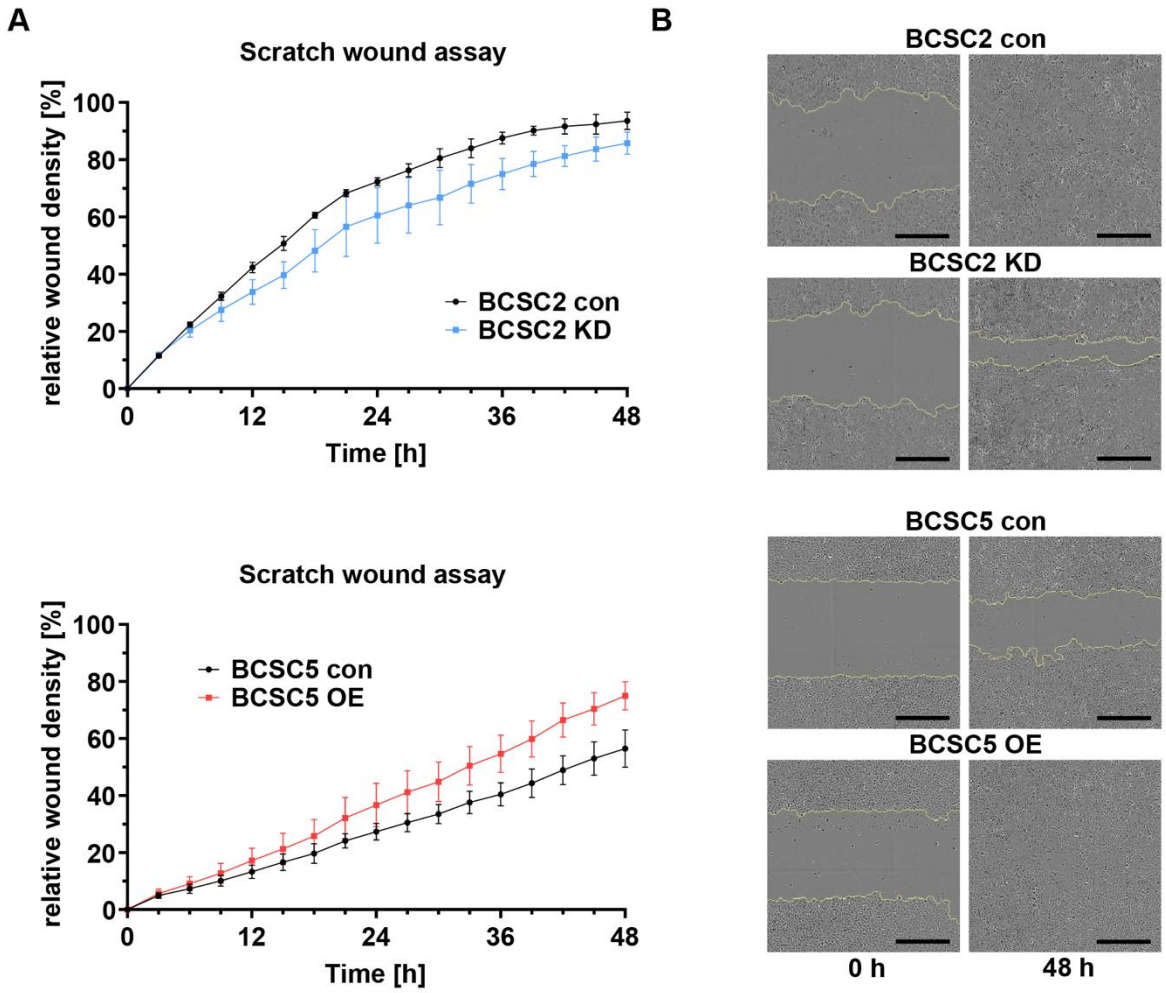


Figure 6. GRHL1 alteration influenced migratory behavior in TNBC cell lines. A, B Scratch wound assays with GRHL1 knockdown (KD) and overexpression (OE) BCSCs in comparison to the control cells (con: scr crispr, stuffer). **A** Quantification of relative wound closure is illustrated as relative wound density in percent. **B** Representative images of TNBC cells after scratch (0h) and after 48h. Detected wound edges are marked by a yellow line. Scale bars represent 400 µm. Data represent mean values ± SD, n = 3.

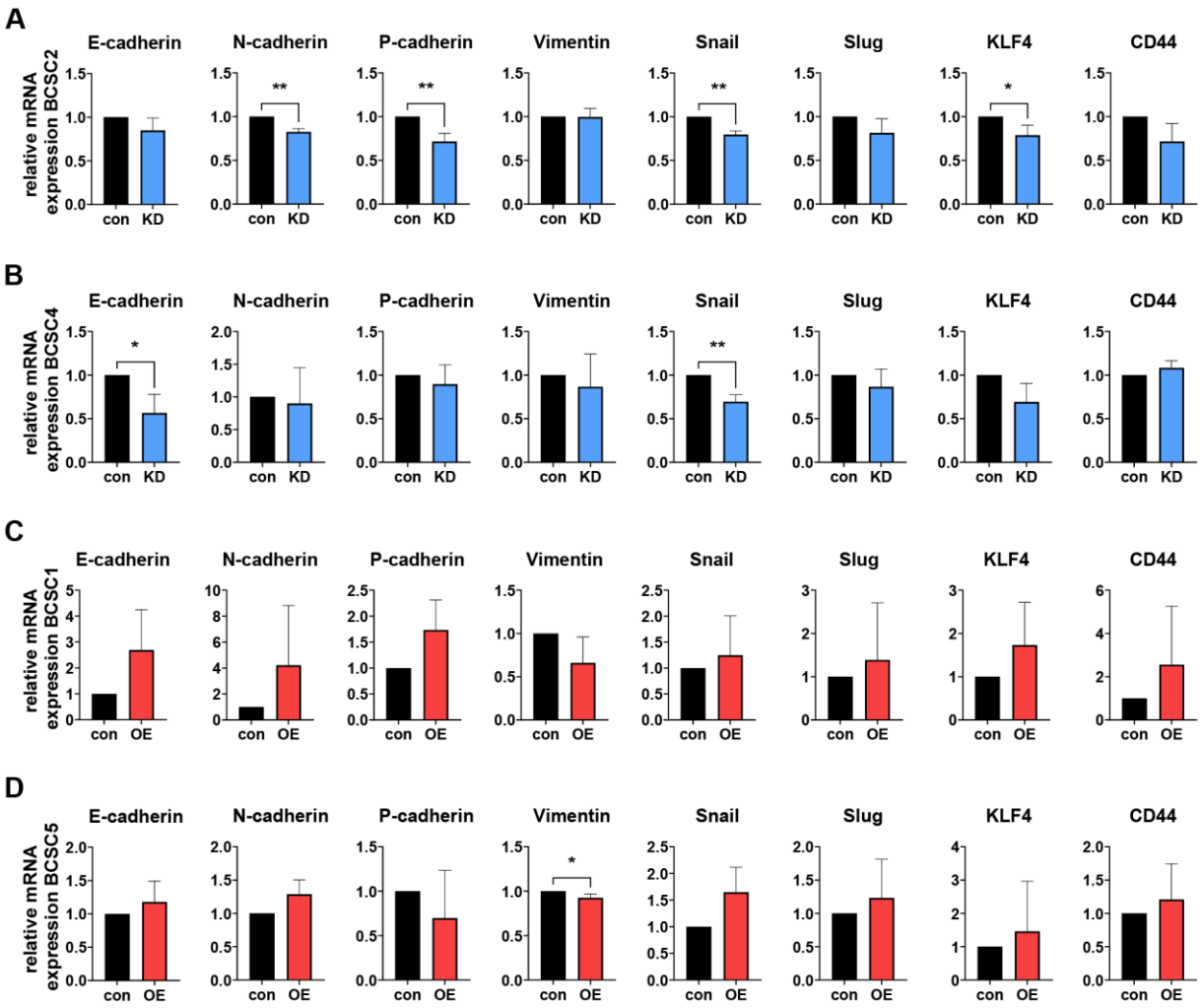


Figure 7. GRHL1 manipulation affects the EMT signature genes of triple-negative BCSCs. A - D RT-qPCR results of EMT-marker (E-cadherin, N-cadherin, P-cadherin, Vimentin, Snail, Slug, KLF4 and CD44) comparing GRHL1 knockdown (KD) and overexpression (OE) to the control (con: crispr scr, stuffer). Cell lines were cultured in 2D (**A**) BCSC2, (**B**) BCSC4, (**C**) BCSC1, (**D**) BCSC5. Mean values ± SD, n = 3. * $p < 0.05$, ** $p < 0.01$ by unpaired t-test.

7. BCSC6 xenograft phenocopies original patient tumor

Previously, we established five triple-negative breast cancer stem cell (BCSC1-5) lines, which all derived from primary invasive breast cancers of individual patients (20, 21). All tumor specimens were collected during surgery after neoadjuvant chemotherapeutic treatment. As presented in Figure 8, we established and characterized BCSC6 with the same methods used in previous studies (20, 21). BCSC6 is a primary breast cancer patient-derived cancer stem cell line of an invasive breast carcinoma of no special type (NST), which was treated neoadjuvant with nab-paclitaxel/carboplatin and epirubicin/cyclophosphamide. The histopathological report characterized the patient tumor as ER 3% positive, PR negative, HER2neu 1+ (low-positive) and Ki67 90% positive, which classifies the tumor as luminal B (26, 27). BCSC6 solely grows non-attached in 3D cell culture conditions, forming spheres of different sizes (Figure 8 A). In a sphere formation assay BCSC6 reached a colony-forming capacity of 78.7% (Figure 8 B). To further characterize the cell line as a BCSC line, stemness features were addressed. We performed a surface marker analysis, which showed 85.5% CD24+/CD44+ and 13.5% CD24-/CD44+ cell phenotype (Figure 8 C) as well as 99,5% EpCAM+/CD49f- cells (Figure 8 D). To verify the stem cell and self-renewing potential of BCSC6 we performed limiting dilution orthotopic xenografts in immunocompromised NOD/SCID mice. BCSC6 demonstrated a high tumorigenic potential as 1000 cells of BCSC6 were able to form tumors (Figure 8 E, F). Extreme Limiting Dilution Analysis (ELDA) calculated the frequency of stem cells to 1/3,007 cells (95% CI: 1/7,863 – 1/1,150) (28). To further demonstrate that the xenograft tumors derived from BCSC6 cells faithfully recapitulate the original patient tumor cytology and cytoarchitecture, we analyzed both tumors immunohistochemically and genetically. Immunohistochemical analysis shows similar expression of ER, PR and HER2 between xenograft and original patient tumor (Figure 8

G). Immunohistochemical staining of PR shows a matching lower positivity in both tumors, but the pathology report classified it as negative (Figure 8 G). Furthermore, immunohistochemical analysis reveals a highly matching expression pattern for the mammary epithelial marker cytokeratin 8, the proliferation marker Ki-67, E-cadherin and Vimentin between BCSC6 xenograft and the corresponding patient tumor (Figure 8 H). RNA sequencing followed by differential gene expression analysis of BCSC6 cell culture, the original patient tumor and the xenograft tumor revealed high degree of similarity between the patient tumor and the corresponding xenograft derived from BCSC6 cells. Especially while exhibiting a more divergent transcriptomic profile in vitro, the BCSC6 cell line retained the ability to generate xenograft tumors that closely recapitulated the transcriptomic profile of the original patient tumor (Figure 8 I). Taken together, these findings demonstrate that very few BCSC6 cells can recapitulate the original patient tumor in histopathological and transcriptomic characteristics as well as biological and functional properties in an in vivo and in vitro model.

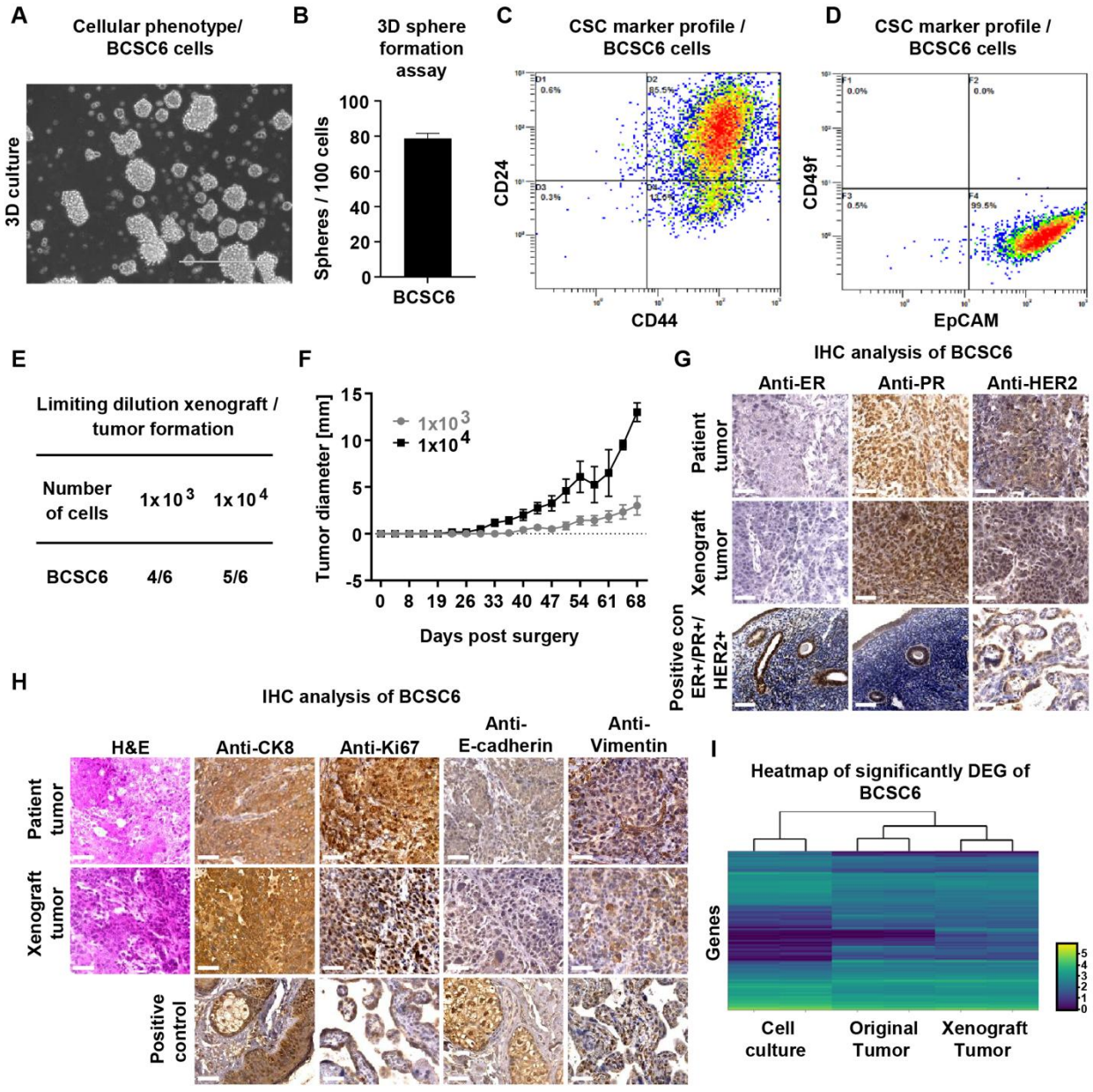


Figure 8. BCSC6 xenograft phenocopies original patient tumor. A Representative image of BCSC6 cells. Scale bar represents 400 µm. **B** 3D sphere formation assay of BCSC6 cells. Data represent mean ± SD. **C, D** Expression pattern of (**C**) CD24, CD44 and (**D**) EpCAM, CD49f in BCSC6 cells analyzed by flow cytometry. **E** Limiting dilution xenograft tumor formation of BCSC6. **F** Growth curves for limiting dilution assay of BCSC6 cell xenografts in immunocompromised mice. Data represent mean + SEM. Initially per

group n = 6, in group with 1 x 10^4 n = 4 from day 54 and n = 2 from day 61 due to predefined endpoint. Data points after euthanasia are therefore not available for these mice. **G**, **H** Immunohistochemical detection of (**G**) ER, PR and HER2 and (**H**) Hematoxylin and eosin (H&E) staining, cytokeratin 8, Ki67, E-cadherin and vimentin on representative sections of the original BCSC6 patient tumor and the BCSC6 xenograft tumor. Scale bar 50 µm. **I** Dendrogram and Heatmap of significantly differentially expressed genes (DEGs) of RNA-sequencing data of BCSC6 cell culture cells, the original tumor samples and xenograft tumor samples derived from BCSC6 cells. Shown are two samples each.

8. GRHL1 manipulation has opposing effects in HER2+ and estrogen receptor positive BCSC in contrast to TNBC

Since the crucial difference in overall survival of patients between low and high GRHL1 expression seems to be related to HER2 as well as hormone receptor expression (Figure 1B - E, Additional File 1 A, B), we investigated the role of GRHL1 in BCSC6. We transduced BCSC6 with GRHL1 KD and OE using lentivirus and validated the KD and OE with RT-qPCR (Figure 9 A) and western blot (Figure 9 B, C). To analyze the effect of GRHL1 on tumorigenicity in BCSC6, we performed sphere formation assays with and without extracellular matrix (ECM). Since BCSC6 usually grows in 3D non-attached spheres, we wanted to assess the effect of GRHL1 in its ability to form spheres in both conditions. KD of BCSC6 only showed a minimal reduction in sphere formation with ECM, while OE increased the ability to form spheres significantly by 42% (Figure 9 D, E). GRHL1 KD of BCSC6 reduced the sphere forming capacity in conditions without ECM by 11%. In comparison to that, OE of BCSC6 without ECM increased the sphere formation

significantly by 12% (Figure 9 F). During analysis we noticed four different phenotypes of spheres. We grouped them into epithelial spheres, where each cell was attached to one another and mesenchymal spheres, where the cells grew more isolated but still grouped together, as well as a more epithelial and more mesenchymal phenotype (Additional File 7). GRHL1 KD of BCSC6 showed more mesenchymal spheres and less epithelial spheres compared to its control. In comparison, GRHL1 OE of BCSC6 revealed a larger proportion of epithelial spheres and a smaller number of mesenchymal spheres (Figure 9 G). OE of GRHL1 in BCSC6 indicate an increased tumorigenicity connected with more epithelial characteristics.

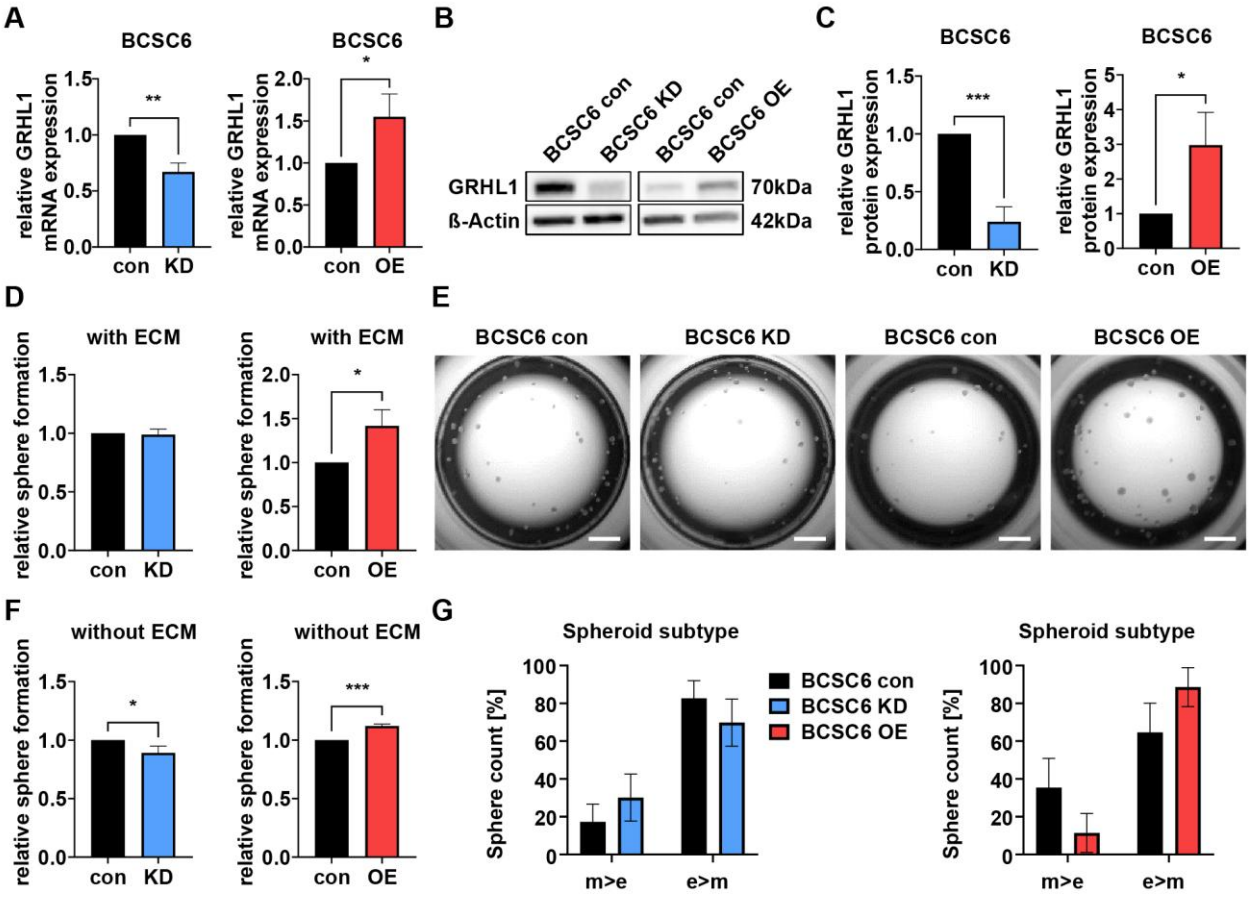


Figure 9. GRHL1 manipulation has opposing effects in luminal B subtype BCSCs in contrast to TNBC. A - C Validation of GRHL1 knockdown (KD) and overexpression (OE) in (**A**) RT-qPCR and (**B, C**) western blot. **B** Representative images of western blot. **C** Quantification of western blot. **D, E, F, G** Sphere formation assays of GRHL1 KD and OE BCSC6 were compared to the control BCSC6 (con: crispr scr, stuffer) using two different methods. **D** Relative sphere formation of BCSC6 GRHL1 KD and OE compared to control. **E** Representative whole well images of the spheres. Scale bars represent 1 mm. **F** Sphere formation of wells with only 1 cell per well were plotted as relative sphere formation of BCSC6 GRHL1 KD and OE compared control. **G** Spheroids from **F** were divided by their phenotype in epithelial, more epithelial, more mesenchymal and mesenchymal spheres. The more epithelial spheroids and epithelial spheroids were plotted together as well as mesenchymal and more mesenchymal spheroids. Representative pictures of those spheres are depicted in Additional File 7. Data represent mean values ± SD, n = 3. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by unpaired t-test.

Discussion

In this study, we investigated the functional role of GRHL1 in the context of triple negative and luminal B breast cancer through functional assays of GRHL1 knockdown (KD) and overexpression (OE) in primary BCSCs and commercial cell lines. GRHL1 OE in TNBC stem cells reduced sphere formation while enhancing migratory behavior and correlated with increased expression of EMT-associated genes. Conversely, GRHL1 KD in the same cells showed reciprocal effects, supporting a functional role in regulation of tumorigenicity and cellular plasticity. GRHL1 OE in luminal B BCSCs presented an

opposite effect with increased sphere formation accompanied by a more epithelial phenotype of spheres, suggesting an interaction between GRHL1 and the expression of hormone receptors as well as HER2.

Amongst grainy head-like genes, GRHL1 was the least investigated in cancer to this date making this the first study of GRHL1 in breast cancer (15). The TCGA data correlation analysis showed a negative correlation of GRHL1 mRNA expression with ER and PR receptor expression. Looking at breast cancer patients in general, low GRHL1 mRNA expression predicts better overall survival with the largest effect in the HER2+ subgroup. However, in the TNBC subgroup, which also has the highest GRHL1 mRNA expression of all subgroups, a high GRHL1 mRNA expression seems to be beneficial for patients' overall survival. These findings suggest that the biological function of GRHL1 is highly dependent on the molecular context, pointing towards a transcriptional regulation coupled with or dependent on hormone signaling and HER2 receptor interaction. Other studies already showed that loss of GRHL2 reduced the expression of androgen receptor and target gene expression through direct manipulation of transcriptional activity in chromatin structures as well as the loss of androgen receptor leading to downregulation of GRHL2 (25, 29, 30). These findings suggest a crucial difference of GRHL1 and GRHL2 function in breast cancer, with GRHL1 having a negative correlation to hormone receptors and GRHL2 increasing the expression of hormone receptors. Different interaction with target genes, a diverse mutation pattern as well as different immunogenic microenvironments could also be a possible explanation for the differences between TNBC and hormone-dependent breast cancer tumors. Since TNBC has the strongest expression of GRHL1 and higher expression in that subgroup correlates with better overall survival, it could potentially serve as a prognostic biomarker.

The lack of an effect on proliferation indicates that GRHL1 is unlikely to regulate cell division. In the context of highly proliferating stem cells in an intermediate state of EMT, this could also be viewed as a stem cell property showing maximal proliferative capacity (31, 32).

KD of GRHL1 in TNBC increased sphere forming capacity and reciprocal GRHL1 OE resulted in a reduction of sphere formation. This implies a reduction of tumor-initiating potential and may reflect alterations in cellular plasticity through GRHL1 modulation in all analyzed TNBC cell lines. This is consistent with the results of Kaplan Meier curves where low GRHL1 mRNA expression correlates with poorer overall survival in the TNBC subgroup. Also, when comparing normal tissue to tumor tissue, GRHL1 is downregulated, which also supports the context of reduced GRHL1 being associated with a more aggressive cellular phenotype. These findings indicate that GRHL1 has a tumor suppressive function in TNBC, which was already shown for neuroblastoma, esophageal squamous cell carcinoma and squamous cell carcinoma of the skin (18, 19, 33).

Higher migratory behavior upon GRHL1 OE was exclusively observed in the triple negative BCSC model and not in commercial TNBC cell lines MDA-MB-468, MDA-MB-231 or the immortalized breast cell line MCF10A. These results indicate a BCSCs-specific effect for the ability to change migratory capacity through GRHL1 while sphere formation was altered across all analyzed cell lines. This BCSC exclusive effect may suggest that the role of GRHL1 in cell motility is dependent on stem cell-associated characteristics like an intermediate state of EMT, which is present in BCSCs but not the commercial lines (22). These observations can be put into clinical context when looking at overall survival data of patients divided into tumor stages. High GRHL1 mRNA expression has an even

greater prognostic relevance in patients with stage III and IV compared to stage I and II tumors showing significantly worse overall survival. These data could support the hypothesis that higher GRHL1 expression led to more migration and thereby metastasis with poorer overall survival in patients with advanced tumor stage. However, the analysis of TCGA data in this case refers to all patients and not only the triple negative subtype, due to a small sample size in this subgroup. Collectively these findings support an oncogenic function of GRHL1 in contrast to the tumor suppressive role when looking at sphere formation in TNBC.

The changes GRHL1 OE and KD had on migration and sphere formation can only be partly explained through the EMT marker analysis. The increase of the EMT master regulator Snail and Slug and the mesenchymal marker N-cadherin in GRHL1 OE as well as the reciprocal decrease in GRHL1 KD may contribute to the enhanced migration and a more mesenchymal phenotype in GRHL1 OE in triple negative BCSCs (31, 34, 35). Consistent with the present findings, recent studies already observed that GRHL1 and GRHL2 positively regulate the epithelial marker E-cadherin (36, 37). But since GRHL1 had no real effect on the expression of vimentin, which is typically upregulated during EMT (38), it creates the image of a complex regulatory environment where the BCSCs are in an intermediate state rather than fully undergoing EMT. Conversely to a real EMT-mediated cadherin switch with creating a highly mesenchymal phenotype GRHL1 OE shows an increased expression of cadherins in general as well as KD of GRHL1 reduced the expression of cadherins. Regarding KLF4 the observed positive regulation with GRHL1 was already shown in recent studies (39). Since KLF4 restores epithelial traits (40), increased KLF4 in GRHL1 OE cells could may explain the higher E-cadherin exhibited in those cells. Looking at BCSCs in comparison to commercial cell lines like the

more epithelial MDA-MB-468 (41) or the very mesenchymal MDA-MB-231 (42), epithelial and mesenchymal markers are highly co-expressed by the same cell line (22). Together this supports the hypothesis that the triple negative BCSCs in their intermediate state of EMT were only slightly shifted to a more mesenchymal phenotype with increased migration and reduced sphere formation. Reciprocal KD of GRHL1 resulted in reduced expression of mesenchymal markers accompanied by a decreased migration and increased sphere formation suggesting a shift towards a more stem cell-associated cellular state.

Recent studies described GRHL1 as a pioneer TF (10), being able to access typically inaccessible chromatin and thereby promote its accessibility and lead to binding of additional TF which can activate target gene transcription (43). This was first described in Grh in Drosophila (44), but similar in studies with MCF7 cells where KD of the human homolog GRHL1-3 showed a reduction in the accessibility of chromatin (45). Being a pioneer transcription factor could explain the ability of GRHL1 to play crucial roles in cell fate decisions and cellular reprogramming by targeting silent chromatin and recruiting co-factors, changing the regulatory network leading to cancer progression (46-48). This may further contribute to the complex regulatory involvement of GRHL1 in stemness and migratory associated pathways.

GRHL1 OE increasing migration while reducing sphere formation in triple negative BCSC shows similarities with previous work by Strietz et al. where treatment of BCSCs with TGF-ß also showed enhanced migration and decreased sphere formation (22). Previous studies observed the relationship between GRHL2 and TGF-ß, focusing on their inverse regulation with GRHL2 suppressing the TGF-ß mediated process of EMT (12, 49). Since

the molecular mechanisms between GRHL1 and TGF-ß remain unknown, this study suggests a functional involvement of GRHL1 in regulating the epithelial-mesenchymal plasticity. These findings raise the possibility that GRHL1, different than GRHL2, may function parallel to TGF-ß modulated pathways by regulating epithelial chromatin accessibility with effects on transcription of epithelial or mesenchymal downstream pathways. This interaction seems like an interesting topic for further investigation and may also elucidate the functional differences among GRHL family members and their distinct biological functions regarding tumorigenicity and cancer progression.

Given the differences in overall survival between TNBC and non-TNBC or especially the HER2+ subtype we characterized a patient-derived primary Luminal B (ER 3% positive, PR negative, HER2neu 1+ (low-positive)) cell line, BCSC6. Interestingly, BCSC6 which solely grows in 3D exhibits a high sphere formation of 78,8% in vitro despite containing only 13,5% CD24-/CD44+ cell population. This suggests that the CD24-/CD44+ cell phenotype may not fully capture the stem cell population in this model and raises the possibility that this marker is less representative for cancer stem cells especially in luminal B breast cancer, which we have also observed in previous studies (20-22). In addition, BCSC6 demonstrated a high tumor-initiation potential. As few as 1000 cells were able to generate tumors in orthotopic xenograft immunocompromised NOD/SCID mice recapitulating the original patient tumor. Therefore, BCSC6 is a suitable model to investigate whether the functional role of GRHL1 differs between breast cancer subtypes.

In contrast to the effects in TNBC, modulation of GRHL1 in the luminal B cell line BCSC6 resulted in opposing effects on sphere formation capacity and the cellular phenotype. Those are consistent with the favorable overall survival for low in comparison to high

GRHL1 mRNA expression. OE of GRHL1 showed increased sphere-formation as well as KD reciprocally reduced the number of spheres under both ECM-containing and ECM-free conditions. Interestingly, characterization of the different spheres revealed a more epithelial phenotype for spheres with an OE of GRHL1 as well as more mesenchymal characteristic spheres for GRHL1 KD. These findings indicate that the functional role of GRHL1 is highly dependent on the molecular context of the breast cancer subtype. Since BCSC6 has only a low positivity for HER2 as well as ER, a different model with a higher positivity of those receptors or even the use of a HER2-enriched phenotype may enhance the presented effects. Also, it is still uncertain whether the effect is driven by hormone receptor or HER2 positivity. Future studies should focus on more cell lines of different breast cancer subtypes to address the underlying subtype-specific functions and involvements of GRHL1 in breast cancer.

This study should be interpreted considering several limitations. Some TCGA subgroups, especially the TNBC subgroup, as well as the patient groups included in the tissue microarray only contain a small sample size, due to the rarity of the disease. This can affect statistical power. Also, in this study GRHL1 was only studied in cell culture and more complex models like xenografts or organoids would shed more light on the tumor microenvironment. Furthermore, the molecular mechanisms behind the observations, including the identification of binding partners, co-factors as well as genomic targets, will require further research. These studies would be essential to explain subgroup specific differences and understand the differences between GRHL family members. Regarding the involvement of GRHL1 on epithelial-mesenchymal plasticity especially the interactions between GRHL1 and EMT transcription factors like Snail, Slug, Twist, ZEB1/2, and the

TGF-ß signaling pathways could provide insights into the molecular function of GRHL1 and could reveal downstream targets.

Conclusion

Collectively, the present findings indicate a dual and context dependent role of GRHL1 in breast cancer. GRHL1 appears to function as tumor suppressor by reducing tumor initiating features as well as it exerts oncogenic effects like increasing migration in triple negative BCSCs. In luminal B breast cancer GRHL1 seems to act completely different with enhanced colony formation accompanied by changes into a more epithelial phenotype upon GRHL1 OE. These functional differences are consistent with subtype-specific prognostic associations observed in Kaplan-Meier overall survival analyses, emphasizing that the biological function of GRHL1 cannot be generalized across all breast cancer subtypes. This is the first study analyzing the functional role of GRHL1 in breast cancer and a correlation of GRHL1 and HER2. To further understand molecular pathways and interactions underlying these subtype-specific mechanisms future research is needed to potentially determine GRHL1 as a prognostic biomarker or therapeutical target.

Materials and Methods

1. Statistical and Survival Analysis of TCGA Data

The TCGA Firehose Legacy Breast Invasive Carcinoma database was accessed through the cBioPortal for Cancer Genomics (50-53) and mRNA expression z-scores

(threshold 2.0) relative to diploid samples (RNA Seq V2 RSEM) data were analyzed. All patients with GRHL1 mRNA expression values from primary tumors were included (n = 1091; n = 1089 with available survival data). The GRHL1 data were matched through their sample ID with the clinical and genomic information of the patient. The data was analyzed using SPSS. GRHL1 mRNA expression data from matched tumor and normal breast tissue samples were downloaded from TCGA Breast Invasive Carcinoma (TCGA-BRCA) project through the National Cancer Institute Genomic Data Commons Data Portal (54, 55). Data were obtained using RNA-sequencing data (Data Category: Transcriptome Profiling, Data Type: Gene Expression Quantification, Workflow Type: STAR-Counts, RNA Sequencing Data: tpm_unstranded). Cases including normal tissue samples (n = 112) were matched through case ID with samples of primary tumor tissue. We excluded cases missing clinical data or without demographics in clinical data (n = 1). Data of matched samples are shown in a Paired Box plot, and statistical significance was analyzed with Wilcoxon test using Python. To analyze the GRHL1 mRNA expression level of different subtypes we downloaded gene expression array data (AgilentG4502A_07_3) from the UCSC Xena platform (56) and matched the samples via sample ID with the molecular subtypes (PAM50_mRNA_nature2012) (57). The data was analyzed using GraphPad Prism 10.0 (GraphPad, San Diego, CA, USA).

2. Tissue Micro Array and Immunohistochemistry

Tissue microarrays (TMAs) were constructed from formalin-fixed, paraffin-embedded tissue specimens obtained from 16 patients with invasive breast cancer. All tissue samples were retrieved from the Institute of Surgical Pathology, University Medical Centre Freiburg, Germany. For each patient, multiple regions of interest (ROIs) were

defined on digitized haematoxylin and eosin (H&E)-stained slides. ROIs were subsequently used for TMA construction. Each TMA contained three tissue cores per patient, with each core measuring 1 mm in diameter. The use of multiple ROIs and multiple cores per patient was intended to partially account for intratumoral heterogeneity. Immuno-histochemical staining for GRHL1 was performed using a polyclonal anti-GRHL1 anti-body (Sigma-Aldrich, HPA005798). The primary antibody was diluted at 1:200 and incubated for 60 minutes. Detection was carried out using an avidin–biotin-based detection system (Dako EnVision FLEX, K800021-2, Dako/Agilent Technologies). Nuclear counterstaining was performed with haemalaun. Staining quality was assessed morphologically. Non-evaluable tissue cores, tissue artefacts and non-representative areas were excluded from further analysis where applicable. The stained slides were digitized using a P1000 whole-slide scanner (3DHISTECH Ltd., Budapest, Hungary). Digital image analysis was performed using HALO AI image analysis software, version 4.2.6399.90 (Indica Labs Inc., Albuquerque, NM, USA). An in-house trained classifier was applied to identify and quantify tumor tissue, stroma and lymphocytes. In addition, an image analysis module trained for nuclear detection was used to quantify immunohistochemical staining intensity, which was classified as absent, low, moderate or strong. The results of the automated tissue classification and staining quantification were visually reviewed to identify obvious segmentation errors or misclassified areas by two pathologists. Areas with relevant artefacts, non-tumor tissue or insufficient tissue quality were excluded from the final analysis where necessary. For each patient, mean and median H-values were calculated from the corresponding tumor cores using RStudio (Version 2026.04.0+526, R Foundation for Statistical Computing, Vienna, Austria). The H-value was used as a quantitative summary measure of GRHL1 immunohistochemical

expression in tumor tissue. Where applicable, H-values were derived from the proportions of tumor nuclei assigned to the respective staining-intensity categories. Statistical analyses were performed using SPSS and R. Overall survival was defined as the time from initial diagnosis or primary surgical treatment to death from any cause or last follow-up, depending on the available clinical follow-up data. GRHL1 expression was evaluated using the cutpointr package (58) to determine optimal cut-points by maximizing the sum of sensitivity and specificity.

Survival analyses were performed using Kaplan–Meier analysis as well as univariable and multivariable Cox proportional hazards regression models. Hazard ratios (HRs) with 95% confidence intervals (CIs) were reported. For multivariable analysis, pathological tumor stage (pT), lymph node stage (pN), tumor grading and patient age were included as covariates. Owing to the limited sample size and the low number of events, the number of covariates was restricted, and selected clinicopathological variables were collapsed to improve model stability. Pathological tumor stage was dichotomized as pT1–2 versus pT3–4. Lymph node stage was categorized into three groups (pN0, pN1 and pN2-3) according to the distribution of available cases and events. For subgroup-specific multivariable analyses, predictors were automatically excluded if only a single category was present, if a predictor level contained no observed events, or if the variable consisted entirely of missing values.

3. BCSC Culture

Breast cancer stem cell (BCSC) lines, mostly triple negative, were previously isolated from patient tumor samples, with patients having undergone chemotherapy

treatment [20-22]. Written informed consent was obtained from all patients prior to sample collection. All procedures involving human specimens were conducted in accordance with the principles of the Declaration of Helsinki. The study protocol and all experimental procedures were approved by the Institutional Review Board of the Independent Ethics Committee of the University of Aachen (ethics vote 206/09). All experiments were performed in accordance with the approved protocol. BCSC lines were cultured in mammary stem cell (MSC) medium consisting of MEBM (Lonza, Basel, Switzerland) supplemented with 2% B27 (Thermo Fisher Scientific, Waltham, MA, USA), 1% Amphotericin B (Sigma-Aldrich, St. Louis, MO, USA), 1% Penicillin-Streptomycin (Gibco), 20 ng/mL EGF (PeproTech, Rocky Hill, NJ, USA), 4 µg/mL heparin (Sigma- Aldrich), 20 ng/mL FGF (PeproTech), 35 µg/mL gentamicin (Thermo Fisher Scientific, Waltham, MA, USA), 500 nM H-1152 (Calbiochem, Sigma-Aldrich, St. Louis, MO, USA) and 0.05% RhoKinase (+R) in 2D (BCSC1-5) or 3D (BCSC6) at 37 °C, 5% CO_2, and 3% O_2. 2D cultures were cultured on tissue culture plastic (TCP) coated with 2% Cultrex (R&D Systems, Minneapolis, MN, USA). Cells were passaged when they reach 80% confluency. Confluent cells were washed with 1x PBS and then detached using Accutase (Innovative Cell Technologies, San Diego, CA, USA) for 15 minutes at 37 °C, 5% CO_2 and 3% O_2. The cell solution was transferred into a falcon and resuspended 1:1 with MSC+R and centrifuged at 200 x g for 3 minutes. The supernatant was aspirated, and the cell pellet was resuspended in 1ml MSC+R. Medium was changed once or twice a week and BCSCs were only cultured until passage 42. To prepare cell pellets, detached cells were washed once with 1x PBS and centrifuged at 200 x g for 3 minutes. The supernatant was aspirated and the cell pellets stored at -80 °C. For cryo-preservation cells were resuspended in 0.5 ml Bambanker (GC Lymphotec Inc., Tokyo, Japan), frozen overnight at -80 °C and stored

in a nitrogen tank afterwards. For thawing, cells were resuspended with 1ml warm MSC+R, centrifuged at 200 x g for 3 min and seeded on coated (except for BCSC6) TCP.

4. Commercial Cell Culture

Commercial cell lines MDA-MB-231, MDA-MB-468 were cultured in DMEM medium (Thermo Fisher Scientific, Waltham, MA, USA) containing 10% FBS (Thermo Fisher Scientific) and 1% Penicillin-Streptomycin (Thermo Fisher Scientific) at 37 °C and 5% CO_2. Immortalized breast cell line MCF10A was cultured in DMEM-F12 medium supplemented with 5% horse serum, 0,2% EGF, 1% hydrocortisone, 0.1% Insulin, 1% hepes serum, 0.01% cholera toxin and 1% Penicillin-Streptomycin (Thermo Fisher Scientific) at 37 °C and 5% CO_2. For subculturing, the cells were first washed with 1x PBS, detached with Trypsin (Thermo Fisher Scientific) and transferred with medium to a 15ml falcon. The cells were centrifuged for 3 minutes at 200 x g, the supernatant discarded, and the cells seeded with medium in an appropriate amount in a new culturing flask. The medium was changed once or twice a week. Cell pellet preparation and cryopreservation were performed as described above for BCSCs. Commercial cell lines were thawed according to the same protocol, using the culture medium mentioned above.

5. Lentiviral Virus preparation and Titration

To produce stable lentiviral plasmids for GRHL1 CRISPR/Cas9 knockdown (GRHL1 KD) and scramble (con), as well as GRHL1 overexpression (OE) and stuffer (con) HEK293T cells were used. Plasmids were designed and purchased from Vectorbuilder

(control CRISPR/Cas9 scrambled vector: VB210702-1063puv, GRHL1 knockdown vector: VB240215-1049rqd, control stuffer vector: VB900122-0484ubz, GRHL1 overexpression vector: VB900013-2878r). 1000 ng of each plasmid was pipetted in the middle of every well of a 6 well plate. Then, 400 µl/well of helper plasmids (700 ng pCMVdR8.74, as well as 350 ng pMDVSVG) in OPTI-MEM (Thermo Fisher Scientific) containing 4 µl Xtreme Gene HP (Roche, Basel, Switzerland) were added. After 30 minutes of incubation at room temperature 2 x 10^6 HEK293T cells resuspended in 1600 µl DMEM were added. The plates were incubated at 37 °C and 5% CO_2. After 24 hours media was replaced by 2 ml/well MSC. After another 24 hours, the viral soup from all wells was harvested, centrifuged at 500 x g for 10 minutes and the supernatant was filtered through a 0.45 µm filter. The plate was cultured with 2 ml/well MSC for the next 24 hours and the process was repeated for harvesting day 2 viral soup. To isolate the virus, media was mixed 4:1 with sucrose buffer, equally distributed into 5ml Eppendorf and centrifuged at 4 °C and 10,000 x g for 4 hours. Supernatant was discarded and the virus was stored at -80 °C. It was titrated on HEK293T cells to define the multiplicity of infection (MOI) 1 of the virus.

6. Lentiviral Transduction

BCSCs for virus transduction were seeded in 12 or 24 well TCP coated with 2% Cultrex (except for BCSC6) and cultured in MSC+R and incubated 24 hours at 37 °C, 5% CO_2 and 3% O_2. Commercial cell lines (MDA-MB-468, MDA-MB-231 and MCF10A) were seeded in 6 or 96 well TCP in DMEM or DMEM F-12 (MCF10A) and incubated 24 hours at 37 °C, 5% CO_2. Medium was replaced by medium containing 15 µg/ml polybrene (Sigma-Aldrich, St. Louis, MO, USA) and virus added in MOI 1 for commercial cell lines,

MOI 5 (BCSC1 and BCSC5 GRHL1 OE) or MOI 10 (BCSC2 and BCSC4 GRHL1 KD, BCSC6 GRHL1 OE and KD). The plate was centrifuged for 1 hour at room temperature and incubated at 37 °C, 5% CO_2 and 3% O_2. After 30 hours medium was changed and after 72 hours selection with medium containing 0.002 mg/ml puromycin was performed. GRHL1 KD and OE were validated through RT-qPCR and western blot.

7. Proliferation Assay

3 x 10^3 cells per well were seeded in a 96 well plate coated with 2% Cultrex (only BCSCs) and cultured at 37 °C, 5% CO_2 and 3% O_2 (only BCSCs) for up to seven days. Confluence over time was measured using IncuCyte Live Cell Analysis System (Sartorius, Göttingen, Germany).

8. Colony Forming Capacity Assay

In a 96 well ultra-low-attachment flat bottom plate 1 x 10^2 (BCSC6) or 2 x 10^2 cells were seeded 1:1 with Cultrex. After 20 minutes in the incubator 150 µl media was added. The spheres were manually counted after 7, 10 or 14 days in the incubator at 37 °C, 5% CO_2 and 3% O_2 (only BCSCs) and imaged with Invitrogen EVOS FL Auto Imaging System (Thermo Fisher Scientific, Waltham, MA, USA) and Zeiss Primovert (Carl Zeiss Microscopy GmbH, Jena, Germany).

9. Single Cell Colony Forming Capacity Assay

One cell of BCSC6 per well was seeded in a 96 well round-bottom ultra-low-attachment plate in 200 µl media and incubated at 37 °C 5% CO_2 and 3% O_2 up to 12 days. Sphere formation starting from only one cell was analyzed with IncuCyte Live-Cell Analysis System.

10. Migration Assay

In a 96 well IncuCyte Image Lock plate coated with 2% Cultrex (only BCSCs) 2 x 10^4 (BCSC2) or 3 x 10^4 cells per well were seeded and cultured for 24 – 48 hours until 100% confluence at 37 °C 5% CO_2 and 3% O_2 (only BCSCs). According to manufacturer protocol a wound was scratched in every well and wound healing was measured every 3 hours over 48 hours with IncuCyte Cell Migration Scratch Wound Analysis. Media was changed every 24 hours.

11. RNA Isolation

MiRNeasy Mini Kit (Qiagen, Hilden, Germany, 217004) was used following manufactures instructions to isolate RNA from cell pellets. RNA concentration was measured using NanoDrop One (Thermo Fisher Scientific, Waltham, MA, USA) and stored by -80 °C.

12. cDNA Synthesis

cDNA synthesis was performed using EvoScript Reverse-Transcriptase Kit (Roche, Basel, Switzerland) for BCSCs and High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Waltham, MA, USA) or iScript™ cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA) for commercial cell lines (MCF10A, MDA-MB-231 and MDA-MB-468) following the manufacturer instructions. For the synthesis in the Eppendorf Mastercycler nexus gradient (Eppendorf SE, Hamburg, Germany) 200 ng or 100 ng of mRNA in 5 µl was used. cDNA was stored at -20 °C.

13. Quantitative Reverse Transcription PCR Analysis

For RT-qPCR the iQ SYBR Green Supermix Kit (Bio-Rad, Hercules, CA, USA) was used according to manufacturer protocol and the RT-qPCR performed using LightCycler480 (Roche, Basel, Switzerland).

Table 3. Primer sequences used for Quantitative Reverse Transcription PCR Analysis

Primer	Forward (5′-3′)	Reverse (5′-3′)
CD44	CAACACAAATGGCTGGTACG	GTGTGGTTGAAATGGTGCTG
E-cadherin	TTCTGGAAGGAATGGAGGAGTC	ACCTGGAATTGGGCAAATGTG
GAPDH	TCGCTCTCTGCTCCTCCTGTTC	CGCCCAATACGACCAAATCC
GRHL1	AGGAGTTTTCTTCCCCTCGG	ACTGTCTCCTGGCTTCTGTC
KLF4	CGAACCCACACAGGTGAGAA	TACGGTAGTGCCTGGTCAGTTC
N-cadherin	GTGCATGAAGGACAGCCTCT	AGCTTCTCACGGCATACACC
P-cadherin	GCTGGGGAAAGTATTCATGG	CCTTTCCTTCAGTGACCTTCTTT
Slug	GGAAGAGATCTGCCAGACGC	CAATGGCAACCAGACAACCG
Snail	CGAGTGGTTCTTCTGCGCTA	TGCAGCTCGCTGTAGTTAGG

Vimentin	CTGCCAACCGGAACAATGAC	GGCATCTACTTCACAGGTGA

14. Protein Isolation and Quantification

Cell pellets were resuspended with lysis buffer (RIPA buffer containing 10% phosphoSTOP (Roche, Basel, Switzerland) and 2% cOmplete mix (Roche, Basel, Switzerland)), incubated for 30 minutes on ice and centrifuged for 15 minutes by 4 °C and 13.000 rpm. The protein concentration of the supernatant was measured with DC Protein Assay Kit II (Bio-Rad, Hercules, CA, USA) and SpectraMax iD3 (Molecular Devices, San Jose, CA, USA) following manufacturer's instructions. Isolated protein was stored by -20 °C.

15. SDS PAGE and Western Blot Analysis

20 µg of isolated protein was diluted in 15 µl with 3,75 µl Lämmli buffer containing 10% 2-Mercaptoethanol and cooked for 5 minutes by 95 °C. Mini-PROTEAN TGX™ Precast Protein Gel (Bio-Rad, Hercules, CA, USA) was loaded with 15 µl protein lysate and Precision Plus Protein Dual Color Standards (Bio-Rad, Hercules, CA, USA) was used as a Marker. The Gel was run with 130 V for 1 up to 1,5 hours followed by 10 minutes of incubation in transfer buffer. For the Transfer of proteins to a PVDF-membrane Trans-Blot Turbo Transfer System (Bio-Rad, Hercules, CA, USA) was used. Membranes were blocked for 1 hour in 5% BSA in TBST and afterwards incubated overnight at 4 °C with the primary antibody diluted in 5% BSA in TBST (GRHL1 1:2000 (Sigma-Aldrich, St. Louis, MO, USA, HPA005798) ß-Actin 1:2000 (Sigma-Aldrich, St. Louis, MO, USA, A5441)). Membranes were washed three times for 5 minutes with TBST before incubating with the

secondary antibody, diluted in 5% BSA in TBST (Polyclonal goat anti-rabbit 1:2000 (Dako, Agilent Technologies, Santa Clara, CA, USA, P044801), Polyclonal rabbit anti-mouse 1:2000 (Dako, Agilent Technologies, Santa Clara, CA, USA, P0260), for 1h at room temperature. For detection of the chemiluminescence signal Miracle Star ECL Reagent (iNtRON Biotechnology, Seongnam, South Korea) und Fusion SL imaging system (Vilber, Collégien, France) was used. Western blots were quantified using ImageJ 15.49.03 (National Institutes of Health, Bethesda, MD, USA).

16. BCSC6 isolation

Primary breast cancer tumors were received from patients during surgery after neoadjuvant chemotherapy treatment as previously described for BCSC1-5 [20-22]. Primary BCSC6 cell line was isolated by mechanical dissociation of the tumor material followed by enzymatic digestion with 6 units of DNAse I (Machery-Nagel, Düren, Germany) and 1 mg Liberase (Roche, Basel, Switzerland) in 5ml DPBS (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) for 1h at 37 °C. After diluting in 10ml DPBS and filtration through a 40 µm cell strainer the cells were centrifuged at 200 x g for 5min, the supernatant was discarded and the cells washed with MEBM (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). Subsequently, 2ml of ACK Lysis-buffer (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) was added to the pellet to remove red blood cells. After 1 minute of incubation time, the suspension was filled up to 6ml with MEBM and centrifuged at 200 x g for 5 minutes. The supernatant was discarded, the cells resuspended in 1ml MEBM and filtered through a 40 µm cell strainer again. After another centrifugation at 200 x g for 5 minutes, the supernatant was discarded and the remaining

cell pellet resuspended in MSC medium. The isolated cells were counted and 2 x 10^4 cells in 200 µl of a 1:1 mixture of MSC medium and ice cold Matrigel (Corning, NY, USA, 354230) were seeded per well in a 24-well low attachment plate (Corning, NY, USA). Once the Matrigel had solidified at 37 °C for 30 min, 500 µL of MSC medium was added to each well. Cells were cultured at 37 °C with low oxygen conditions (3% O_2, 5% CO_2, buffered with N_2). Following stable proliferation in 3D culture, cells were expanded under 2D culture conditions. BCSC6 was isolated in 2021 and authenticated by the high-throughput Multiplex human cell authentication test (MCA) developed at the DKFZ in 2016 (59). Every 3 months mycoplasma tests were performed via PCR detection utilizing the positive control provided free of charge by the Leibniz-Institute DSMZ. All experiments with BCSCs were performed using cells in a passage lower than 42.

17. Flow Cytometry

To analyze marker of stemness BCSC6 was cultured in a 60 mm dish. Cells were detached using Accutase for 25 minutes, resuspended in MSC+R medium and counted as mentioned before. 300K cells of BCSC6 were aliquoted into each Fluorescence-Activated Cell Sorting (FACS) tube and washed with PBS and centrifuged at 1000rpm for 5 minutes two times. A third wash was performed using FACS buffer (2% FBS in PBS) and centrifuged again. After discarding the supernatant, the antibodies anti-CD24 (eBioscience, Thermo Fisher Scientific, Waltham, MA, USA, APC, 17-0247-42, 1:40), anti-CD44 (eBioscience; PE, 12-0441-82, 1:40) anti-EpCAM (eBioscience; eFlour 660, 50-9326-42, 1:40), anti-CD49f (eBioscience; eFlour 710, 46-0495-82; 1:40) were diluted in 100 µl of FACS buffer and added to the cells. After 20 minutes of incubation at room

temperature without light cells were washed two times using and resuspended in 700 µl FACS buffer. FACS sort was performed with Beckman Coulter Life Sciences Navios (Beckman Coulter Life Sciences, Brea, CA, USA) with 3 lasers analyzed with Navios Software.

18. Orthotopic Breast cancer Xenograft

In vivo experiments were conducted in accordance with German Animal Welfare regulations and approved by Institutional Animal Care and Use Committee as described in animal protocol G13/114. Shortly, female mice were group-housed (maximum five per cage) in climate-controlled, individually ventilated cages (IVCs) on a 12-hour light/dark cycle with nesting material provided for environmental enrichment and constant access to food and water. Upon arrival, the animals underwent a seven-day acclimation period and were managed exclusively using low-stress, non-aversive handling techniques.

1 x 10^3 and 1 x 10^4 BCSC6 cells were mixed with 1 x 10^6 irradiated (15 Gy) Hs27 fibroblasts (ATCC, Manassas, VA, USA). Cells were mixed with Matrigel at a 1:1 ratio and MSC+R to a total volume of 40 µl and put on ice. Cell mixtures were transplanted into the mammary fat pads of non-obese diabetic/severe combined immunodeficiency (NOD/SCID) female mice in isoflurane inhalator-based anesthesia. Therefore, a small sagittal incision (approx. 1 cm) was made on the shaved and sterilized abdomen to provide bilateral access to the mammary glands. For each mouse, two injections were performed, one on each side and the surgical incisions sutured with 8/0 thread. The animals were sacrificed as the xenografts reached a diameter close to 15 mm. Three mice, with two mammary fat pads each, were utilized per cell dilution (1 x 10^3, 1 x 10^4 BCSC6

cells, six animals in total). Animals were monitored each day measuring weight and twice per week measuring tumor growth using caliper. Due to tumor size, in the group with 1 x 10^4 cells one animal was euthanized on day 54 (afterwards n = 4, 2 animals) and one on day 61 (n = 2, 1 animal). Limiting dilutions were calculated using the Extreme Limiting Dilution Analysis software (28). Tumor growth curves are displayed in Graph Pad Prism 10.0 (GraphPad, San Diego, CA, USA).

19. Immunohistochemistry and Haematoxylin and Eosin staining

Patient tumor tissue specimens as well as xenograft samples were fixed in 4% PFA at 4 °C for 24h, embedded in paraffin, sliced into 3-4 µm thick sections and mounted onto glass slides (VWR SuperFrost Plus Microscopic Slides, 631-9483). Slides were incubated at 37 °C overnight immediately after cutting and subsequently deparaffinized using xylol and washed with ethanol in descending concentrations, then rinsed with distilled water. For heat-induced antigen retrieval, slides were immersed in citrate buffer at pH 6 (Dako, Agilent Technologies, Santa Clara, CA, USA, 169984-2) and placed in microwave cuvettes. The slides were heated in a microwave at 800 W for 30 seconds and subsequently incubated in a water bath for 1h at 97 °C. After 15 minutes of cool down the slides were washed with distilled water for 5 minutes. Endogenous peroxidase was blocked using 5% methanol-based hydrogen peroxide solution (H_2O_2, VWR, 23615261) for 15 minutes and washed with distilled water another time. Then the slides were placed in TBST buffer (Cell Signaling Technology, Danvers, MA, USA, 9997 S, 10 x TBST diluted 1:10 in distillated H_2O). The tissue slides were stained with primary antibodies in required dilution using Dako Antibody Diluent (Dako Agilent, S-302283-2) and incubated in a wet

channel at 4 °C overnight. The following primary antibodies were used: anti-ER (Santa Cruz Biotechnology, Inc., Dallas, TX, USA, sc-543), anti-PR (Santa Cruz Biotechnology, sc-538), anti-HER2 (Santa Cruz Biotechnology, sc-264), anti-cytokeratin 8 (Novus Biologicals, Centennial, CO, USA, NBP2-16094), anti-Ki67 (Sigma-Aldrich, St. Louis, MO, USA, SAB4501880), anti-E-cadherin (Sigma-Aldrich, SAB4503751), anti-Vimentin (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA, PA5-27231). The following day, the slides were washed with TBST buffer and incubated with the secondary antibody (polyclonal goat anti-rabbit Antibody-HRP, 1:100 dilution in Dako Agilent, P044801) for 45 minutes at 37 °C. After washing the slides in TBST buffer, they were incubated for 9 minutes in 3,3'-Diaminobenzidin (DAB) solution (Dako Agilent Kit, K346811-2) in the dark. Slides were subsequently rinsed with mineral water and then distilled water until counter-stained with hematoxylin (Dako Agilent, CS70030-2) for 5 minutes and washed with mineral water and distilled water again. The tissue slides were immersed in increasing concentrations of ethanol, followed by 5 minutes of incubation in xylol before being covered with mounting medium Entellan (Sigma-Aldrich, St. Louis, MO, USA, 1079610500) and coverslip.

For Haematoxylin and Eosin staining (H&E) paraffin sections (3–4 µm) were deparaffinized and rehydrated as described before. Sections were stained with hematoxylin for 5 min, blued in running tap water for 10 min, rinsed in distilled water, and counterstained with eosin for 10 min. After staining sections were rinsed in distilled water, dehydrated, cleared in xylene and mounted as mentioned before.

Slides were scanned by the Institute of Pathology, University Hospital Aachen, using NanoZoomer 2.0 HT (Hamamatsu Photonics K.K., Hamamatsu, Japan). The resulting

digital whole-slide images were analyzed using QuPath 0.7.0. (University of Edinburgh, Edinburgh, UK). Tissue samples of uterus, placenta and skin with sebaceous glands were used as positive control for ER, PR (uterus), HER2, Ki67, vimentin (placenta) and cytokeratin 8 and E-cadherin (skin) staining.

20. Heatmap DEGS

Frozen tumor and patient tumor tissue were pulverized under liquid nitrogen to obtain a fine powder and transferred into pre-cooled microcentrifuge tubes. RNA of tumor tissue samples as well as BCSC6 cells was isolated as described above. RNA concentration was measured using NanoDrop One and RNA samples frozen at -80 °C.

3´mRNA library preparation and sequencing as well as bioinformatic analysis was performed by the Genomics Facility, a core facility of the Interdisciplinary Center for Clinical Research (IZKF) Aachen within the Faculty of Medicine at RWTH Aachen University. Total RNA integrity and concentration were assessed using a TapeStation (Agilent Technologies, Santa Clara, CA, USA) and a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), respectively. For each sample, 500 ng of total RNA was used as input for library preparation. Sequencing libraries were generated using the Lexogen QuantSeq 3' mRNA-Seq Library Prep Kit (Lexogen GmbH, Vienna, Austria) FWD with UDIs, according to the manufacturer's instructions. Following library construction, library concentration and fragment size distribution were reassessed using Qubit and TapeStation, respectively, to confirm library quality prior to sequencing. Sequencing was performed on a NextSeq 500/550 platform using a Mid Output v2.5 Kit (150 cycles) (Illumina, San Diego, CA, USA). Libraries were pooled and loaded at a final concentration

of 1.6 pM, spiked with 1% PhiX as an internal control, and sequenced in single-end mode with a read length of 75 bases. FASTQ files were generated using bcl2fastq (Illumina) and processed with nf-core/rnaseq version 3.12.0 (60), executed in Nextflow using Docker containers (61). Reads were quality-controlled with FastQC, trimmed with Trim Galore, and aligned to the human GRCh38 primary assembly using STAR version 2.7.10a with GENCODE release 44 annotation (62, 63). STAR parameters were adapted for 3′-mRNA sequencing, including clipping of terminal poly(A) sequences. Transcript abundance was estimated using Salmon version 1.10.1 with transcript-length correction disabled (64), and transcript-level estimates were summarized to gene-level counts using tximport version 1.30.0 without abundance or length scaling (65). Differential expression analysis was performed in R version 4.3.3 using DESeq2 version 1.42.1 (66). Xenograft and original tumor samples were analyzed using a paired design that included tumor identity as a blocking factor, whereas comparisons involving BCSC6 samples were analyzed as unpaired. Genes with a Benjamini–Hochberg-adjusted P-value below 0.05 were considered differentially expressed.

21. Statistical Analysis

Experiments were performed in triplicates, if not mentioned differently. All data were shown as mean ± SD. TCGA data as well as Kaplan-Meier-Analysis of tissue microarray was analyzed using IBM SPSS 29.0.2.0 (SPSS, Chicago, IL, USA). GraphPad Prism 10.0 (GraphPad, San Diego, CA, USA) was used for statistical analysis of all experimental data. Associations between clinicopathological parameters and GRHL1 mRNA and protein expression levels were assessed using Pearson´s chi-square test (two-sided) and

Spearman's rank correlation coefficients. Overall survival (OS) was analyzed using Kaplan–Meier survival curves and log-rank tests. Datasets of two groups were compared using two tailed unpaired t-test. Ordinary one-way ANOVA with Holm- Šídák correction was used for data comparing more than two groups. Data was considered as statistically significant when p-value was less than 0.05. Defined p-value: * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$; **** = $p < 0.0001$.

Abbreviations

The following abbreviations are used in this manuscript:

BCSC	Breast cancer stem cells
ECM	Extracellular matrix
EMP	Epithelial-mesenchymal plasticity
EMT	Epithelial-mesenchymal transition
H&E	Haematoxylin and eosin
KD	Knockdown
MET	Mesenchymal-epithelial transition
MOI	Multiplicity of infection
MSC	Mammary Stem Cell
MSC+R	Mammary Stem Cell + Rho kinase
OE	Overexpression

ROIs	Regions of interest
RT-qPCR	Reverse transcription quantitative PCR
TCGA	The Cancer Genome Atlas
TCP	Tissue culture plates
TMAs	Tissue microarrays
TNBC	Triple-negative breast cancer
TF	Transcription factor

Supplementary Information

Additional File 1. Patients with higher stage breast cancer show better overall survival with low GRHL1 mRNA expression.

Additional File 2. GRHL1 expression in commercial cell lines and validation of GRHL1 knockdown and overexpression.

Additional File 3. Knockdown and overexpression of GRHL1 does not affect proliferation in MDA-MB-468 and MCF10A cell line.

Additional File 4. GRHL1 knockdown and overexpression does not change the morphology of TNBC cell lines.

Additional File 5. GRHL1 knockdown increases sphere forming capacity while overexpression decreases sphere formation.

Additional File 6. Knockdown and overexpression of GRHL1 do not affect migration in MDA-MB-468 and MCF10A cells.

Additional File 7. Morphology of BCSC6 spheres grown from 1 cell in anchorage-independent cell culture

Declarations

Ethics approval and consent to participate

Written informed consent was obtained from all patients prior to sample collection. All experiments involving human specimens were conducted in accordance with the principles of the Declaration of Helsinki. The study protocol and all experimental procedures were approved by the Institutional Review Board of the Independent Ethics Committee of the University of Aachen (ethics vote 206/09, 08.08.2018) and performed in accordance with it and all relevant guidelines and regulations. The maximum tumor size permitted by the ethics committee was 15 mm and was not exceeded in any animal during the study.

Consent for publication:

Not applicable.

Availability of data and materials

The mRNA-sequencing datasets analyzed in this study are accessible through the TCGA datasets, the TCGA Firehose Breast Invasive Carcinoma database on cBioPortal for Cancer Genomics and the Xenabrowser dataset. The RNA-seq data of BCSC6 shown in

this study are publicly available in the NCBI Gene Omnibus (GEO) under accession number GSE351107.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Deutsche Krebshilfe (#70115581), the DFG (RTG 2375) and the BMBF (#100450773) to J. Maurer.

Authors´ contributions

Conceptualization and methodology: JM, ESp, MR

TCGA data analysis and interpretation: ESp, MR

Tissue microarray analysis and survival analysis: ESc, PB

Breast cancer stem cell (BCSC) functional assays, cell culture and experiments: ESp

Cell culture and experiments using commercial breast cancer cell lines: CD

BCSC6 characterization: HH, TH

BCSC6 xenograft transplantation and animal experiment: TH, JM

Data analysis: ESp

Interpretation of data: ESp, JM

Writing – original draft: ESp

Writing – review and editing: JM, LH, TH, ESp, ESt

Funding acquisition: JM, ESt

Supervision: JM, ESt

All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the Flow Cytometry Facility, a core facility of the Interdisciplinary Center for Clinical Research (IZKF) Aachen within the Faculty of Medicine at RWTH Aachen University and by the German Research Foundation (DFG), project ID 439895892. This work was further supported by the Genomics Facility of the Interdisciplinary Center for Clinical Research (IZKF) within the Faculty of Medicine at RWTH Aachen University. We also thank the staff of the animal facility at UKA for continuous support as well as Catharina Maaßen, Birgit Klein and Anja Steinle for their technical support.

Authors´ information

Emma Spreemann espreemann@ukaachen.de

Celasty Das cdas@uni-bonn.de

Elisabeth Schygulla Elisabeth.schygulla@uniklinik-freiburg.de

Thomas Hansen mail@thomas-hansen.eu

Hannah Heidenreich hannah.heidenreich@rwth-aachen.de

Lena Hahnfeld lhahnfeld@ukaachen.de

Michael Rose mrose@ukaachen.de

Peter Bronsert peter.bronsert@uniklinik-freiburg.de

Elmar Stickeler estickeler@ukaachen.de

Jochen Maurer jmaurer@ukaachen.de

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209-49.
2. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. N Engl J Med. 2010;363(20):1938-48.
3. Hua Z, White J, Zhou J. Cancer stem cells in TNBC. Semin Cancer Biol. 2022;82:26-34.
4. Giatromanolaki A, Sivridis E, Fiska A, Koukourakis MI. The CD44+/CD24- phenotype relates to 'triple-negative' state and unfavorable prognosis in breast cancer patients. Med Oncol. 2011;28(3):745-52.
5. Luo M, Brooks M, Wicha MS. Epithelial-mesenchymal plasticity of breast cancer stem cells: implications for metastasis and therapeutic resistance. Curr Pharm Des. 2015;21(10):1301-10.
6. Liao TT, Yang MH. Revisiting epithelial-mesenchymal transition in cancer metastasis: the connection between epithelial plasticity and stemness. Mol Oncol. 2017;11(7):792-804.
7. Park M, Kim D, Ko S, Kim A, Mo K, Yoon H. Breast Cancer Metastasis: Mechanisms and Therapeutic Implications. Int J Mol Sci. 2022;23(12).
8. Ting SB, Caddy J, Hislop N, Wilanowski T, Auden A, Zhao LL, et al. A homolog of Drosophila grainy head is essential for epidermal integrity in mice. Science. 2005;308(5720):411-3.
9. Mathiyalagan N, Miles LB, Anderson PJ, Wilanowski T, Grills BL, McDonald SJ, et al. Meta-Analysis of Grainyhead-Like Dependent Transcriptional Networks: A Roadmap for Identifying Novel Conserved Genetic Pathways. Genes (Basel). 2019;10(11).
10. Sundararajan V, Pang QY, Choolani M, Huang RY. Spotlight on the Granules (Grainyhead-Like Proteins) - From an Evolutionary Conserved Controller of Epithelial Trait to Pioneering the Chromatin Landscape. Front Mol Biosci. 2020;7:213.
11. Lammert FC, Pannhausen J, Noetzel E, Friedland F, Wirtz J, Herfs Y, et al. Dual role of GRHL3 in bladder carcinogenesis depending on histological subtypes. Mol Oncol. 2024;18(6):1397-416.
12. Cieply B, Riley Pt, Pifer PM, Widmeyer J, Addison JB, Ivanov AV, et al. Suppression of the epithelial-mesenchymal transition by Grainyhead-like-2. Cancer Res. 2012;72(9):2440-53.

13. Mlacki M, Kikulska A, Krzywinska E, Pawlak M, Wilanowski T. Recent discoveries concerning the involvement of transcription factors from the Grainyhead-like family in cancer. Exp Biol Med (Maywood). 2015;240(11):1396-401.
14. Zhao P, Guo S, Tu Z, Di L, Zha X, Zhou H, et al. Grhl3 induces human epithelial tumor cell migration and invasion via downregulation of E-cadherin. Acta Biochim Biophys Sin (Shanghai). 2016;48(3):266-74.
15. Gasperoni JG, Fuller JN, Darido C, Wilanowski T, Dworkin S. Grainyhead-like (Grhl) Target Genes in Development and Cancer. Int J Mol Sci. 2022;23(5).
16. He Y, Gan M, Wang Y, Huang T, Wang J, Han T, et al. EGFR-ERK induced activation of GRHL1 promotes cell cycle progression by up-regulating cell cycle related genes in lung cancer. Cell Death Dis. 2021;12(5):430.
17. Guo S, Bai W, Cui F, Chen X, Fang X, Shen H, et al. Exploration of the Correlation Between GRHL1 Expression and Tumor Microenvironment in Endometrial Cancer and Immunotherapy. Pharmgenomics Pers Med. 2024;17:91-103.
18. Fabian J, Lodrini M, Oehme I, Schier MC, Thole TM, Hielscher T, et al. GRHL1 acts as tumor suppressor in neuroblastoma and is negatively regulated by MYCN and HDAC3. Cancer Res. 2014;74(9):2604-16.
19. Mlacki M, Darido C, Jane SM, Wilanowski T. Loss of Grainy head-like 1 is associated with disruption of the epidermal barrier and squamous cell carcinoma of the skin. PLoS One. 2014;9(2):e89247.
20. Metzger E, Stepputtis SS, Strietz J, Preca BT, Urban S, Willmann D, et al. KDM4 Inhibition Targets Breast Cancer Stem-like Cells. Cancer Res. 2017;77(21):5900-12.
21. Li X, Strietz J, Bleilevens A, Stickeler E, Maurer J. Chemotherapeutic Stress Influences Epithelial-Mesenchymal Transition and Stemness in Cancer Stem Cells of Triple-Negative Breast Cancer. Int J Mol Sci. 2020;21(2).
22. Strietz J, Stepputtis SS, Follo M, Bronsert P, Stickeler E, Maurer J. Human Primary Breast Cancer Stem Cells Are Characterized by Epithelial-Mesenchymal Plasticity. Int J Mol Sci. 2021;22(4).
23. Yuan M, Wang J, Fang F. Grainyhead-Like Genes Family May Act as Novel Biomarkers in Colon Cancer. Onco Targets Ther. 2020;13:3237-45.
24. Burks HE, Pokorny JL, Koetsier JL, Roth-Carter QR, Arnette CR, Gerami P, et al. Melanoma cells repress Desmoglein 1 in keratinocytes to promote tumor cell migration. J Cell Biol. 2023;222(11).
25. Reese RM, Harrison MM, Alarid ET. Grainyhead-like Protein 2: The Emerging Role in Hormone-Dependent Cancers and Epigenetics. Endocrinology. 2019;160(5):1275-88.
26. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol. 2011;22(8):1736-47.
27. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol. 2013;24(9):2206-23.
28. Hu Y, Smyth GK. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J Immunol Methods. 2009;347(1-2):70-8.

29. Wang Z, Coban B, Wu H, Chouaref J, Daxinger L, Paulsen MT, et al. GRHL2-controlled gene expression networks in luminal breast cancer. Cell Commun Signal. 2023;21(1):15.
30. Paltoglou S, Das R, Townley SL, Hickey TE, Tarulli GA, Coutinho I, et al. Novel Androgen Receptor Coregulator GRHL2 Exerts Both Oncogenic and Antimetastatic Functions in Prostate Cancer. Cancer Res. 2017;77(13):3417-30.
31. Brabletz S, Schuhwerk H, Brabletz T, Stemmler MP. Dynamic EMT: a multi-tool for tumor progression. Embo j. 2021;40(18):e108647.
32. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell. 2008;133(4):704-15.
33. Li M, Li Z, Guan X, Qin Y. Suppressor gene GRHL1 is associated with prognosis in patients with oesophageal squamous cell carcinoma. Oncol Lett. 2019;17(5):4313-20.
34. Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in cancer. Nat Rev Cancer. 2018;18(2):128-34.
35. Lee JY, Kong G. Roles and epigenetic regulation of epithelial-mesenchymal transition and its transcription factors in cancer initiation and progression. Cell Mol Life Sci. 2016;73(24):4643-60.
36. Werth M, Walentin K, Aue A, Schönheit J, Wuebken A, Pode-Shakked N, et al. The transcription factor grainyhead-like 2 regulates the molecular composition of the epithelial apical junctional complex. Development. 2010;137(22):3835-45.
37. Kohn KW, Zeeberg BM, Reinhold WC, Pommier Y. Gene expression correlations in human cancer cell lines define molecular interaction networks for epithelial phenotype. PLoS One. 2014;9(6):e99269.
38. Usman S, Waseem NH, Nguyen TKN, Mohsin S, Jamal A, Teh MT, et al. Vimentin Is at the Heart of Epithelial Mesenchymal Transition (EMT) Mediated Metastasis. Cancers (Basel). 2021;13(19).
39. Kotarba G, Taracha-Wisniewska A, Miller M, Dabrowski M, Wilanowski T. Transcription factors Kruppel-like factor 4 and paired box 5 regulate the expression of the Grainyhead-like genes. PLoS One. 2021;16(9):e0257977.
40. Yori JL, Seachrist DD, Johnson E, Lozada KL, Abdul-Karim FW, Chodosh LA, et al. Krüppel-like factor 4 inhibits tumorigenic progression and metastasis in a mouse model of breast cancer. Neoplasia. 2011;13(7):601-10.
41. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell. 2006;10(6):515-27.
42. Riehl BD, Kim E, Lee JS, Duan B, Yang R, Donahue HJ, et al. The Role of Fluid Shear and Metastatic Potential in Breast Cancer Cell Migration. J Biomech Eng. 2020;142(10).
43. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 2011;25(21):2227-41.
44. Nevil M, Bondra ER, Schulz KN, Kaplan T, Harrison MM. Stable Binding of the Conserved Transcription Factor Grainy Head to its Target Genes Throughout Drosophila melanogaster Development. Genetics. 2017;205(2):605-20.
45. Jacobs J, Atkins M, Davie K, Imrichova H, Romanelli L, Christiaens V, et al. The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. Nat Genet. 2018;50(7):1011-20.

46. Kuzmich AI, Tyulkina DV, Vinogradova TV, Sverdlov ED. [Pioneer Transcription Factors in Normal Development and in Carcinogenesis]. Bioorg Khim. 2015;41(6):636-43.
47. Iwafuchi-Doi M, Zaret KS. Cell fate control by pioneer transcription factors. Development. 2016;143(11):1833-7.
48. Nevil M, Gibson TJ, Bartolutti C, Iyengar A, Harrison MM. Establishment of chromatin accessibility by the conserved transcription factor Grainy head is developmentally regulated. Development. 2020;147(5).
49. Chen W, Kang KL, Alshaikh A, Varma S, Lin YL, Shin KH, et al. Grainyhead-like 2 (GRHL2) knockout abolishes oral cancer development through reciprocal regulation of the MAP kinase and TGF-β signaling pathways. Oncogenesis. 2018;7(5):38.
50. TCGA Firehose Legacy data in cBioPortal for Cancer Genomics. https://www.cbioportal.org/ Accessed 19 December 2024.
51. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401-4.
52. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6(269):pl1.
53. de Bruijn I, Kundra R, Mastrogiacomo B, Tran TN, Sikina L, Mazor T, et al. Analysis and Visualization of Longitudinal Genomic and Clinical Data from the AACR Project GENIE Biopharma Collaborative in cBioPortal. Cancer Res. 2023;83(23):3861-7.
54. Genomic Data Commons Data Portal of the National Cancer Institute. https://portal.gdc.cancer.gov Accessed 21 April 2024.
55. Heath AP, Ferretti V, Agrawal S, An M, Angelakos JC, Arya R, et al. The NCI Genomic Data Commons. Nature Genetics. 2021;53(3):257-62.
56. UCSC Xena. https://xenabrowser.net/ Accessed 20 January 2025.
57. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38(6):675-8.
58. Thiele C, Hirschfeld G. cutpointr: Improved Estimation and Validation of Optimal Cutpoints in R. Journal of Statistical Software. 2021;98(11):1-27.
59. Castro F, Dirks WG, Fähnrich S, Hotz-Wagenblatt A, Pawlita M, Schmitt M. High-throughput SNP-based authentication of human cell lines. Int J Cancer. 2013;132(2):308-14.
60. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol. 2020;38(3):276-8.
61. Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. Nat Biotechnol. 2017;35(4):316-9.
62. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
63. Frankish A, Carbonell-Sala S, Diekhans M, Jungreis I, Loveland JE, Mudge JM, et al. GENCODE: reference annotation for the human and mouse genomes in 2023. Nucleic Acids Res. 2023;51(D1):D942-d9.
64. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417-9.

65. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 2015;4:1521.
66. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

Supplementary Files

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

- AdditionalFile17final.docx
- WesternBlotsUncropped.pdf