## Supplementary Materials

# Heparan sulfate 3-O-sulfotransferase 3B1 (HS3ST3B1) is associated with invasive and mesenchymal-like phenotype in breast cancer cells and promotes chemo-resistance through ac-tivation of PDGF-Rβ pathway

Agnès Denys [1], Mariama Diawara [1,2], Clémence Messmer [1,3], Emilie Cadart [1], and Fabrice Allain [1,*]

[1] Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France; agnes.denys@univ-lille.fr (A.D.); emilie.cadart@univ-lille.fr (E.C.); fabrice.allain@univ-lille.fr (F.A.)

[2] Biology Department, Université de Moncton, Moncton, New-Brunswick, Canada, E1A 3E9; emd2061@umoncton.ca (M.D.)

[3] Laboratory for Cell Signaling and Cancer, Maisonneuve-Rosemont Hospital Research Center, Montréal, Québec, Canada, QC H1T 2M4; clemence.messmer@umontreal.ca (C.M.)

* Correspondence: fabrice.allain@univ-lille.fr; Tel.: +33-(0)3-20-43-40-10

Figure S1. Overexpression of *HS3ST3B1* mRNA correlates with mesenchymal-like phenotype in BrCa cell lines and patients.

Figure S2. Comparison of the phosphorylation status of RTK in parental *versus* HS3ST3B1-overexpressing MDA-MB-231 cells.

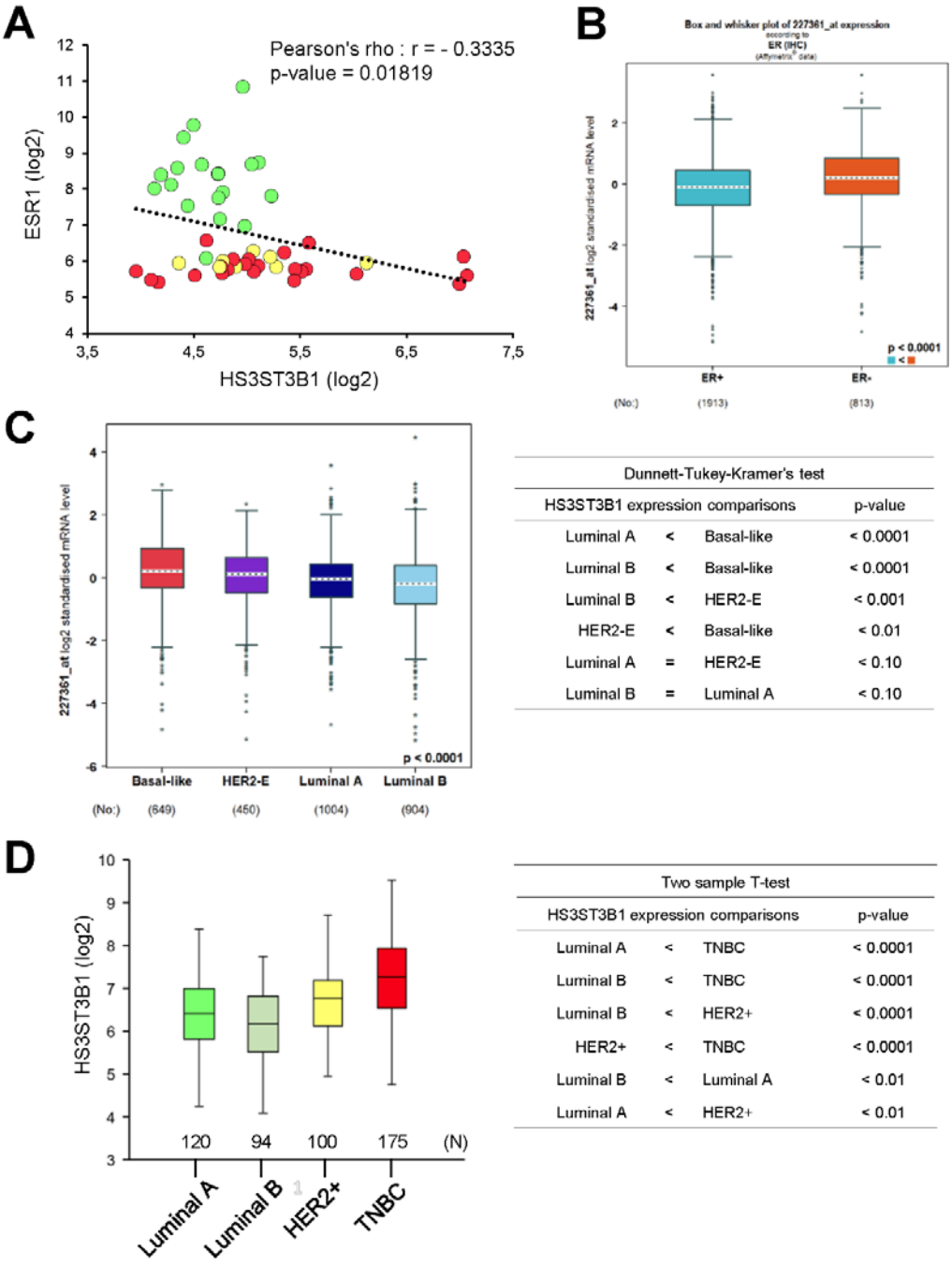
Figure S3. Comparison of the phosphorylation status of the main signaling molecules in parental *versus* HS3ST3B1-overexpressing MDA-MB-231 cells.

Figure S4. Association of high *HS3ST3B1* expression with time survival in BrCa patients.

Table S1. Analysis of the correlations of *HS3ST3B1* mRNA with the transcripts expression levels of the genes encoding EMT regulators, mesenchymal markers and pro-invasive MMP.

Table S2. Analysis of the correlations of *HS3ST3B1* mRNA with the transcripts expression levels of the genes encoding DDR2, EphA1, EphA10, Insulin-R, PDGF-Rβ and ROR1.

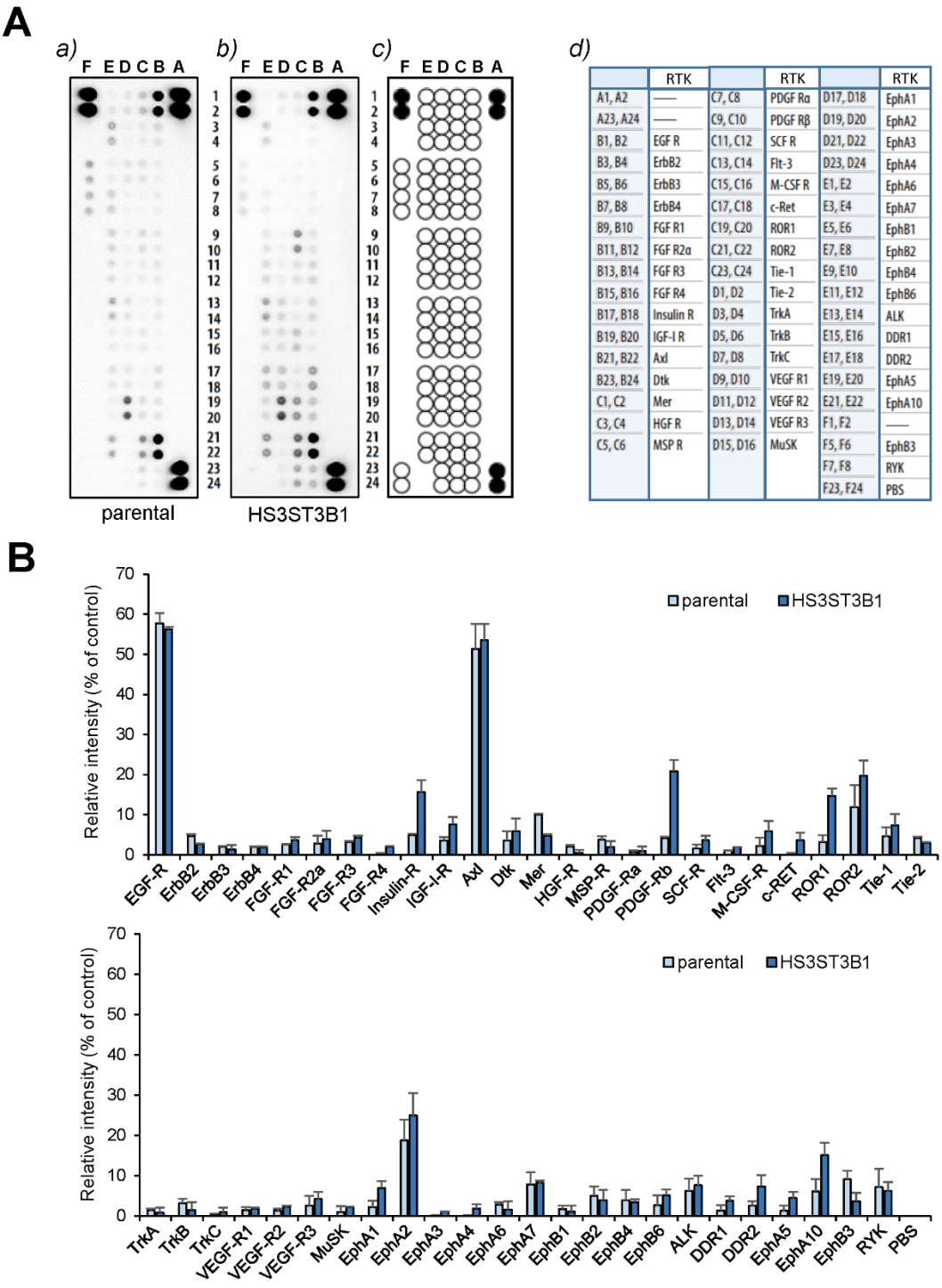
Table S3. Sets of primers used for RT-qPCR analysis.




| Dunnett-Tukey-Kramer's test | | | |
|---|---|---|---|
| HS3ST3B1 expression comparisons | | | p-value |
| Luminal A | < | Basal-like | < 0.0001 |
| Luminal B | < | Basal-like | < 0.0001 |
| Luminal B | < | HER2-E | < 0.001 |
| HER2-E | < | Basal-like | < 0.01 |
| Luminal A | = | HER2-E | < 0.10 |
| Luminal B | = | Luminal A | < 0.10 |

| Two sample T-test | | | |
|---|---|---|---|
| HS3ST3B1 expression comparisons | | | p-value |
| Luminal A | < | TNBC | < 0.0001 |
| Luminal B | < | TNBC | < 0.0001 |
| Luminal B | < | HER2+ | < 0.0001 |
| HER2+ | < | TNBC | < 0.0001 |
| Luminal B | < | Luminal A | < 0.01 |
| Luminal A | < | HER2+ | < 0.01 |

**Figure S1.** Overexpression of *HS3ST3B1* mRNA correlates with mesenchymal-like phenotype in BrCa cell lines and patients. **(A)** Analysis of the correlation between the expression of *HS3ST3B1* and *ESR1* transcripts. Data were obtained from the collection of BrCa Cell Lines [27] in UCSC-Xena data portal. BrCa cell lines were categorized according to the luminal (N = 19) (green dots), HER2+ (N = 8) (yellow dots) and TNBC (N = 22) (red dots) subtypes. **(B)** Comparison of the expression of *HS3ST3B1* in ER+ (N = 1913) versus ER- (N = 813) BrCa patients (P < 0.01). **(C)** Analysis of the expression of *HS3ST3B1* in luminal A (N = 1004), luminal B (N = 904), HER2+ (N = 450) and basal-like (N = 649) subtypes. In both analysis, data were obtained from the bc-GenExMiner v4.5 portal (http://bcgenex.ico.unicancer.fr). The screening condition was the Affymetrix Jetset probe ID # "227361_at". **(D)** Analysis of the expression of *HS3ST3B1* mRNA in luminal A (N = 120), luminal B (N = 94), HER2+ (N = 100) and TNBC (N = 175) subtypes, using data form GENT-2 portal (http://gent2.appex.kr).




| | RTK | | RTK | | RTK |
|---|---|---|---|---|---|
| A1, A2 | —— | C7, C8 | PDGF Rα | D17, D18 | EphA1 |
| A23, A24 | —— | C9, C10 | PDGF Rβ | D19, D20 | EphA2 |
| B1, B2 | EGF R | C11, C12 | SCF R | D21, D22 | EphA3 |
| B3, B4 | ErbB2 | C13, C14 | Flt-3 | D23, D24 | EphA4 |
| B5, B6 | ErbB3 | C15, C16 | M-CSF R | E1, E2 | EphA6 |
| B7, B8 | ErbB4 | C17, C18 | c-Ret | E3, E4 | EphA7 |
| B9, B10 | FGF R1 | C19, C20 | ROR1 | E5, E6 | EphB1 |
| B11, B12 | FGF R2α | C21, C22 | ROR2 | E7, E8 | EphB2 |
| B13, B14 | FGF R3 | C23, C24 | Tie-1 | E9, E10 | EphB4 |
| B15, B16 | FGF R4 | D1, D2 | Tie-2 | E11, E12 | EphB6 |
| B17, B18 | Insulin R | D3, D4 | TrkA | E13, E14 | ALK |
| B19, B20 | IGF-I R | D5, D6 | TrkB | E15, E16 | DDR1 |
| B21, B22 | Axl | D7, D8 | TrkC | E17, E18 | DDR2 |
| B23, B24 | Dtk | D9, D10 | VEGF R1 | E19, E20 | EphA5 |
| C1, C2 | Mer | D11, D12 | VEGF R2 | E21, E22 | EphA10 |
| C3, C4 | HGF R | D13, D14 | VEGF R3 | F1, F2 | —— |
| C5, C6 | MSP R | D15, D16 | MuSK | F5, F6 | EphB3 |
| | | | | F7, F8 | RYK |
| | | | | F23, F24 | PBS |

**Figure S2.** Comparison of the phosphorylation status of RTK in parental *versus* HS3ST3B1-overexpressing MDA-MB-231 cells. Following serum-starvation for 3 h, cells were stimulated for 15 min in the presence of complete culture medium containing 10% of FCS. The arrays were then incubated with 300 µg of total protein extracts and immune-stained with a peroxidase-conjugated phospho-tyrosine detection antibody. **(A)** Representative images of phospho-RTK arrays (panel a, parental cells; panel b, HS3ST3B1-overexpressing cells). The array coordinates (panel c) and the list of the 49 RTK (panel d) are given on the right side of the figure. Black dots represent phospho-tyrosine positive controls: A1-A2; A23-A24, G1-G2. **(B)** Histograms represent the densitometric quantification of the phosphorylation status of each RTK. Data were normalized relative to the positive control dots. The results were obtained from three experiments performed independently.

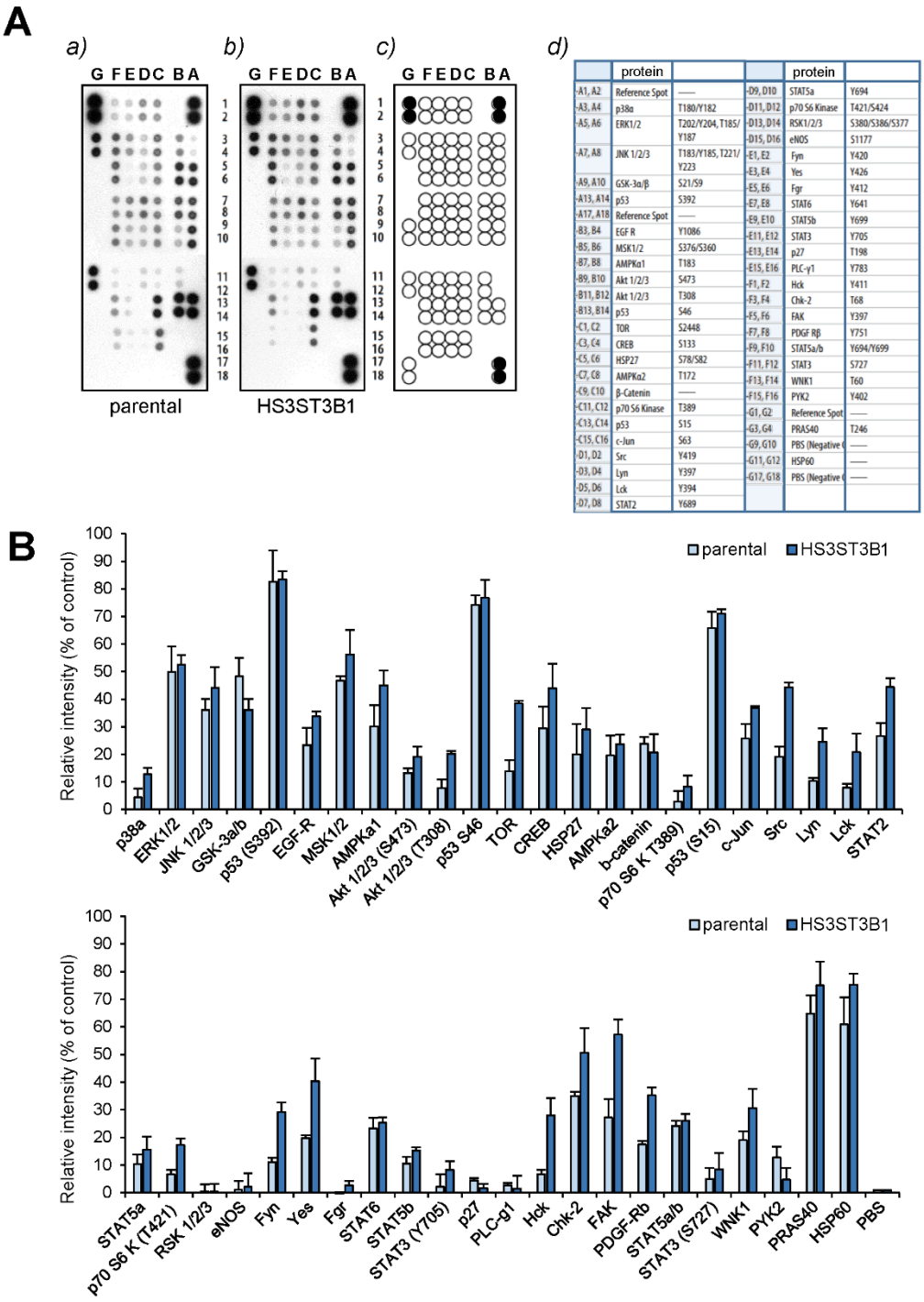



**Figure S3.** Comparison of the phosphorylation status of the main signaling molecules in parental *versus* HS3ST3B1-overexpressing MDA-MB-231 cells. Following serum-starvation for 3 h, cells were stimulated for 15 min in the presence of complete culture medium containing 10% of FCS. The arrays were then incubated with 300 µg of total protein extracts and immune-stained with the detection antibody cocktails. **(A)** Representative images of kinase arrays (panel a, parental cells; panel b, HS3ST3B1-overexpressing cells). The array coordinates (panel c) and the list of signaling molecules (panel d) are given on the right side of the figure. Black dots represent phospho-tyrosine positive controls: A1-A2; A17-A18, F1-F2. **(B)** Histograms represent the densitometric quantification of the phosphorylation status of each signaling molecules. Data were normalized relative to the positive control dots. The results were obtained from three experiments performed independently.

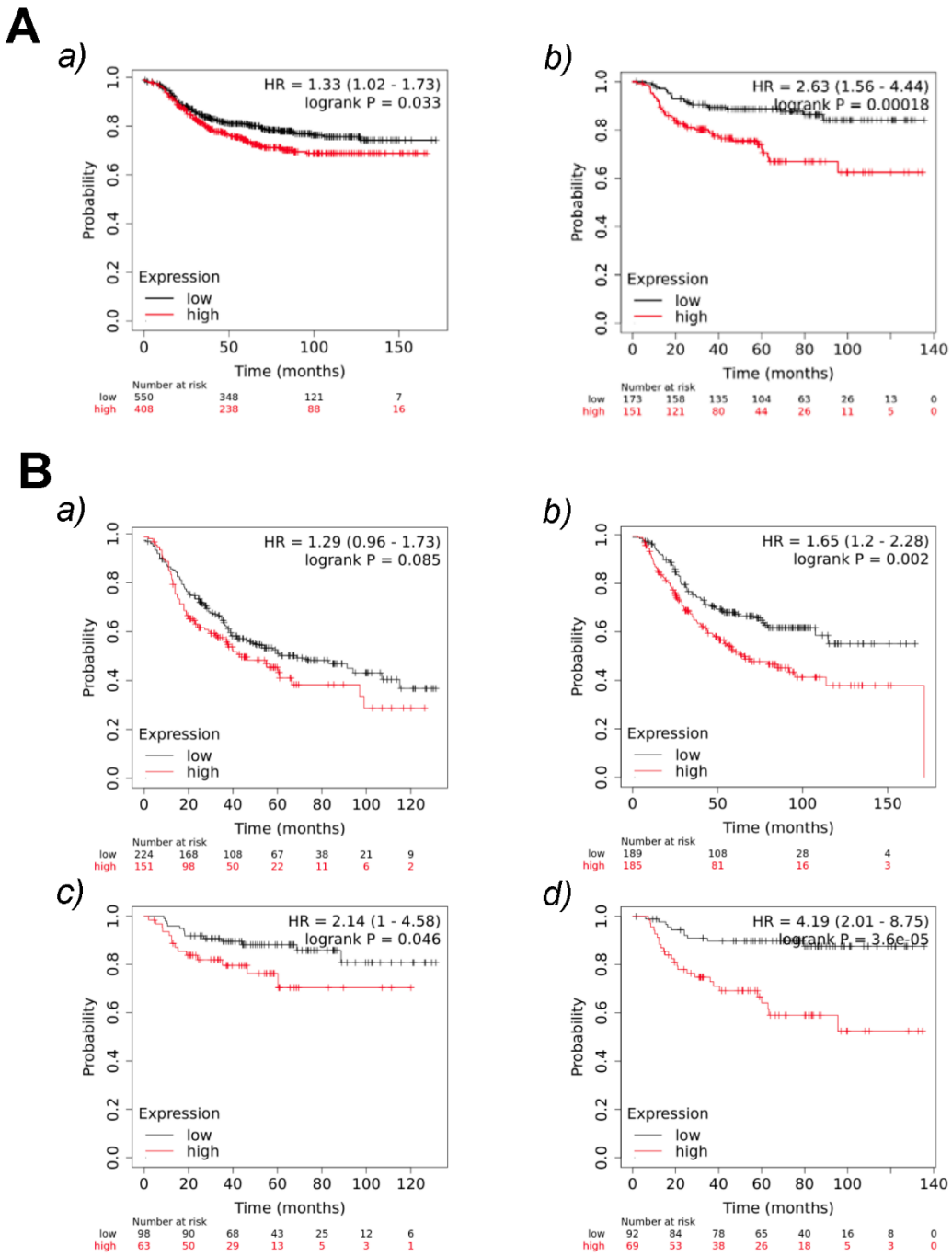



**Figure S4.** Association of high *HS3ST3B1* expression with time survival in BrCa patients. **(A)** Kaplan–Meier analysis of the relationships between *HS3ST3B1* mRNA expression and DMFS in patients with BrCa. Panel a: not restricted analysis (N = 958); panel b: restricted analysis to patients with chemotherapy (N = 324). **(B)** BrCa patients with chemotherapy were filtered by median expression of *PDGFRB*. The correlations of *HS3ST3B1* expression with RFS (panel a: low *PDGFRB* level subgroup; panel b, high *PDGFRB* level subgroup, N = 749) and DMFS (panel c: low *PDGFRB* level subgroup; panel d, high *PDGFRB* level subgroup, N =322) were analyzed in the different patient subgroups. Kaplan-Meier plots were generated through the Jetset probes #"227361_at" and #"202273_at" for *HS3ST3B1* and *PDGFRB* transcripts, respectively, using the KM plotter software (https://www.kmplot.com).

**Table S1.** Analysis of the correlations of *HS3ST3B1* mRNA with the transcripts expression levels of the genes encoding EMT regulators, mesenchymal markers and pro-invasive MMP. Data were obtained from the collection of BrCa Cell Lines [27] and the TCGA BrCa cohorts (UCSC-Xena data portal). The Pearson correlation coefficient above 0.3 was considered as positive correlation.

| | BrCa cell lines ( N = 54) | | BrCa tumor samples ( N = 1247) | |
|---|---|---|---|---|
| | Pearson's rho | p-value | Pearson's rho | p-value |
| *SNAI1* | -0.09522 | 0.4935 | 0.1702 | 2.276e-9 |
| *SNAI2* | 0.4863 | 0.0001964 | 0.3153 | 1.611e-29 |
| *TWIST1* | 0.4238 | 0.001422 | 0.2438 | 6.080e-18 |
| *CDH2* | 0.2777 | 0.04211 | 0.2220 | 4.539e-15 |
| *FN1* | 0.1614 | 0.2437 | 0.3165 | 9.823e-30 |
| *VIM* | 0.3562 | 0.008254 | 0.2275 | 9.139e-16 |
| *MMP2* | 0.3697 | 0.005979 | 0.3171 | 7.432e-30 |
| *MMP9* | 0.01576 | 0.9099 | 0.1899 | 2.353e-11 |
| *MMP14* | 0.3242 | 0.01683 | 0.2926 | 1.777e-25 |

**Table S2.** Analysis of the correlations of *HS3ST3B1* mRNA with the transcripts expression levels of the genes encoding DDR2, EphA1, EphA10, Insulin-R, PDGF-Rβ and ROR1. Data were obtained from the collection of BrCa Cell Lines [27] and the TCGA BrCa cohorts (UCSC-Xena data portal). The Pearson correlation coefficient above 0.3 was considered as positive correlation.

| | BrCa cell lines ( N = 54) | | BrCa tumor samples ( N = 1247) | |
|---|---|---|---|---|
| | Pearson's rho | p-value | Pearson's rho | p-value |
| *DDR2* | 0.1333 | 0.3368 | 0.4396 | 1.046e-58 |
| *EPHA1* | 0.2715 | 0.04714 | - 0.0322 | 0.2614 |
| *EPHA10* | 0.1288 | 0.3535 | - 0.04861 | 0.08992 |
| *INSR* | 0.1422 | 0.3051 | - 0.02232 | 0.4365 |
| *PDGFRB* | 0.4035 | 0.002508 | 0.3608 | 9.344e-39 |
| *ROR1* | 0.2086 | 0.1302 | 0.4818 | 8.557e-72 |

**Table S3.** Sets of primers used for RT-qPCR analysis. Synthetic primers were designed by using Primer-Blast (NCBI) according to the published DNA sequences.

| | | |
|---|---|---|
| *HS3ST3B1* | 5'-ctggcttcaggcaaggagat-3' (forward)<br>5'-gacccactgaaaaagctgga-3' (reverse) | NM_006041.3 |
| *MMP2* | 5'-ctcatcgcagatgcctggaa-3' (forward)<br>5'-ttcaggtaataggcacccttgaaga-3'(reverse) | NM_004530.6 |
| *MMP9* | 5'-acgcacgacgtcttccagta-3' (forward)<br>5'-ccacctggttcaactcactcc-3' (reverse) | NM_004994.3 |
| *MMP14* | 5'-tgcctaccgacaagattgatg-3'(forward)<br>5'-atcccttcccagactttgatg-3'(reverse) | NM_004995.4 |
| *PDGFRB* | 5'-ccaatgagggtgacaacga-3' (forward)<br>5'-ggccctcgtcagcaacct-3' (reverse) | NM_002609.4 |
| *HPRT* | 5'-gaccagtcaacaggggacat-3 (forward)<br>5'-aacacttcgtggggtccttttc-3' (reverse) | NM_000194.3 |