Published in final edited form as:
Cancer Cell. 2024 January 08; 42(1): 52–69.e7. doi:10.1016/j.ccell.2023.11.008.

B3GALT6 Promotes Dormant Breast Cancer Cell Survival and Recurrence by Enabling Heparan Sulfate-Mediated FGF Signaling

Amulya Sreekumar[1,2], **Michelle Lu**[1,2], **Biswa Choudhury**[4], **Tien-chi Pan**[1,2], **Dhruv K. Pant**[1,2], **Matthew R. Lawrence-Paul**[1,2], **Christopher J. Sterner**[1,2], **George K. Belka**[1,2], **Takashi Toriumi**[1,2], **Brian A. Benz**[1,2], **Matias Escobar-Aguirre**[1,2], **Francesco E. Marino**[1,2], **Jeffrey D. Esko**[4], **Lewis A. Chodosh**[1,2,3,*]

[1]Department of Cancer Biology, Perelman School of Medicine at the University of Pennsylvania, PA 19104, USA.

[2]Abramson Family Cancer Research Institute, Perelman School of Medicine at the University of Pennsylvania, PA 19104, USA.

[3]Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, PA 19104, USA.

[4]Department of Cellular and Molecular Medicine, Glycobiology Research and Training Center, University of California, San Diego, La Jolla, CA 92093, USA.

Summary

Breast cancer mortality results from incurable recurrences thought to be seeded by dormant, therapy-refractory residual tumor cells (RTCs). Understanding the mechanisms enabling RTC survival is therefore essential for improving patient outcomes. Here, we derive a dormancy-associated RTC signature that mirrors the transcriptional response to neoadjuvant therapy in patients and is enriched for extracellular matrix-related pathways. In vivo CRISPR-Cas9 screening of dormancy-associated candidate genes identifies the galactosyltransferase B3GALT6 as a functional regulator of RTC fitness. B3GALT6 is required for glycosaminoglycan (GAG) linkage to proteins to generate proteoglycans, and its germline loss-of-function in patients causes skeletal dysplasias. We find that B3GALT6-mediated biosynthesis of heparan sulfate GAGs predicts poor patient outcomes, promotes tumor recurrence by enhancing dormant RTC survival in

---

[*]**Lead contact:** Lewis A. Chodosh, M.D., Ph.D., Correspondence: chodosh@pennmedicine.upenn.edu Room 614 BRB II/III, 421 Curie Boulevard, Philadelphia, PA 19104-6160, USA.

Author contributions
Conceptualization, A.S., J.D.E., and L.A.C.; Methodology, A.S., M.L., B.C., T.P., D.K.P., M.R.L-P., C.J.S., G.K.B., T.T., B.A.B., M.E., and F.E.M.; Investigation, A.S., M.L., B.C., T.P., D.K.P., and M.R.L-P.; Writing, A.S. and L.A.C.; Funding Acquisition, A.S. and L.A.C.

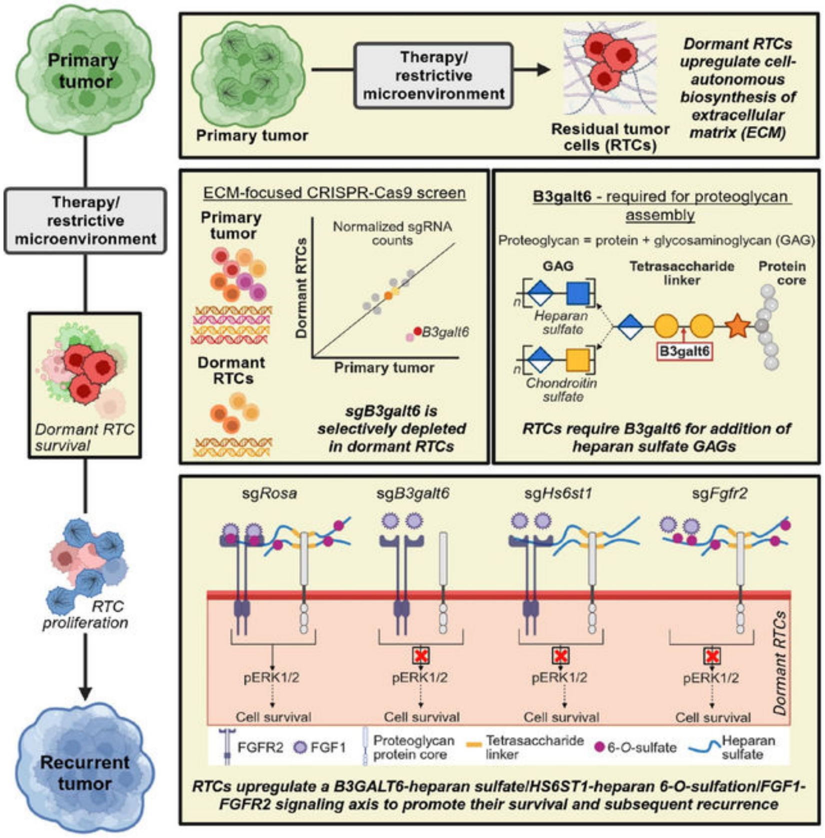
Declaration of interests
L.A.C. has served as an expert consultant to Teva Pharmaceuticals, Eisai, Sanofi, Eli Lilly, Whittaker, Clark and Daniels, Wyeth, Imerys, Colgate, Becton Dickinson, Sterigenics, and the U.S. Department of Justice in litigation.

multiple contexts, and does so via a B3GALT6-heparan sulfate/HS6ST1-heparan 6-*O*-sulfation/FGF1-FGFR2 signaling axis. These findings implicate B3GALT6 in cancer and nominate FGFR2 inhibition as a promising approach to eradicate dormant RTCs and prevent recurrence.

eTOC blurb

Sreekumar *et al.* identify a previously uncharacterized function for the galactosyltransferase B3galt6 in cancer, specifically enabling dormant residual tumor cell survival following therapy or in a restrictive microenvironment by promoting heparan sulfate-mediated FGF signaling. These findings implicate a poorly studied class of molecules, glycans, in breast cancer dormancy and recurrence.

Graphical Abstract



Keywords

B3GALT6; dormancy; breast cancer; proteoglycans; glycosaminoglycans; glycans; heparan sulfate; HS6ST1; 6-*O*-sulfation; FGFR2

Introduction

Despite advances in early detection and treatment, breast cancer remains the leading cause of cancer-related deaths among women worldwide[1]. Mortality results predominantly from incurable recurrences that arise years, or even decades, following treatment of the primary tumor (PT)[2,3]. Since recurrent tumors arise from residual tumor cells (RTCs) that survive therapy and may reside in a reversibly quiescent, non-proliferative state of cellular dormancy[4], depleting this critical pool of cells by targeting their survival mechanisms represents an attractive approach to preventing breast cancer recurrence.

Dormancy has been observed in RTCs that survive therapy or that encounter a foreign microenvironment following their dissemination from a PT[5]. Experimental models recapitulating these paradigms of cellular dormancy include i) therapy-associated models that mimic RTC survival following targeted therapy by downregulating or inhibiting oncogenes such as *Her2*[6,7], *Wnt1*[7,8], and *Fgfr1*[9] in genetically engineered mouse (GEM) models, and ii) models that recapitulate interactions between RTCs and a foreign microenvironment, such as the D2.OR-D2A1 paired cell line model[10–12].

Recent data suggest marked similarities in dormant RTC gene expression profiles that are independent of both the stimulus responsible for dormancy entry and the location of RTCs at local or distant metastatic sites[7,13]. Consequently, investigating conserved, tumor cell-autonomous elements of dormancy-associated gene expression signatures may provide a tractable approach for identifying unique dependencies of dormant RTCs that could be targeted to induce their elimination, thereby preventing recurrence.

Results

Dormant tumor cells display cell-autonomous upregulation of ECM-related genes following therapy

To identify mechanisms underlying dormant RTC survival and persistence, we developed and characterized doxycycline-inducible genetically engineered mouse (GEM) models of breast cancer[6–8,14–19]. These models of cellular dormancy[7] enable robust spatiotemporal regulation of oncogenic signaling that parallel effects of targeted therapy, thereby permitting study of the role of dormancy in spontaneous recurrence.

Dormant RTCs that survive PT regression induced by Her2 downregulation in *MMTV-rtTA;TetO-Her2/neu* (*MTB/TAN*) mice[6], or Wnt1 downregulation in *MMTV-rtTA;TetO-Wnt1* (*MTB/TWNT*) mice[8], were previously isolated to generate gene expression signatures[7]. Dormancy signatures from *MTB/TAN* and *MTB/TWNT* RTCs were highly concordant, suggesting conserved elements of the dormant state[7].

Dormancy-associated changes in gene expression in RTCs in vivo following therapy could be imposed by the microenvironment or by cell-autonomous determinants. To identify cell-autonomous regulators of RTC survival and dormancy, we cultured PT cells (PTCs) from *MTB/TAN* mice in vitro (Fig. 1A). Maintaining Her2-dependent PTCs in the presence of doxycycline induced Her2 expression and resulted in high levels of cell proliferation at D0 (*baseline*), as indicated by Ki67 expression and EdU incorporation (Fig. 1B, C). Withdrawing doxycycline (Her2 *deinduction*) elicited the rapid decay of Her2 expression and proliferation, with negligible levels observed at D7, D14, and D28, as well as increased cell death. Importantly, re-addition of doxycycline to surviving RTCs at each of these time points (D7+, D14+, D28+; Her2 *reinduction*) rapidly restored Her2 levels and proliferation to baseline levels (Fig. 1B, C). Thus, RTC quiescence is reversible, which is a *sine qua non* of cellular dormancy, confirming that this system recapitulates key features of dormancy observed in vivo[7,14,20].

To identify a core set of dormancy-associated genes, we performed RNA-sequencing on samples at *baseline*, *deinduction*, and *reinduction* time points. We evaluated whether gene expression changes in vitro (D0 vs. D28) are conserved with those occurring in RTCs in vivo (PTCs vs. D28 RTCs)[7]. This revealed extensive overlap among downregulated genes (hypergeometric p=7.20e-18) (Fig. S1A). As anticipated, gene ontology (GO) analysis of the overlapping set of 471 downregulated genes identified marked enrichment for pathways related to cell cycle, cellular biosynthesis, and translation (Fig. S1B).

We also found a highly significant overlap among genes upregulated during dormancy in vitro (D0 vs. D28) and in vivo (PTCs vs. D28 RTCs) (309 genes, hypergeometric p=3.40e-60) (Fig. 1D). Intriguingly, GO analysis of this overlapping gene set identified enrichment for multiple pathways related to the organization and metabolism of ECM components (Fig. 1E). We refer to the overlapping in vivo and in vitro-derived up- and down-regulated 780 gene set as the core RTC signature.

A core RTC dormancy signature recapitulates neoadjuvant therapy-associated gene expression changes and predicts favorable outcomes in patients

To evaluate whether the core RTC signature exhibited by dormant tumor cells following Her2 downregulation is clinically relevant, we interrogated data from six datasets of paired gene expression profiles for primary breast cancers and residual tumors following neoadjuvant therapy in patients[21–27]. Five of these six datasets displayed significant enrichment for the core RTC signature post-therapy (Fig. 1F), even after excluding proliferation-associated genes (Fig. S1C). Notably, enrichment for this signature was independent of both the type (chemotherapy/endocrine therapy) and duration of therapy. These findings indicate that therapy-refractory tumor cells that persist in patients exhibit changes in gene expression similar to those associated with cellular dormancy in mice.

To evaluate the prognostic power of the core RTC signature, we performed a meta-analysis of recurrence data derived from patients with breast cancer. We anticipated that the core RTC signature might include genes whose expression was associated with better patient outcome (*e.g.*, pro-cell cycle arrest), as well as genes associated with worse patient outcome (*e.g.*, pro-RTC survival). Analogous to our prior observations employing an in vivo RTC dormancy signature[7], we found that early-stage patients whose PTs were enriched for the core RTC signature exhibited a striking decrease in recurrence risk (HR=0.09; p=3.8e-24), ostensibly reflecting an increased propensity of such tumors to display indolent properties (Fig. S1D). This indicates that the core RTC signature is weighted towards genes associated with better patient outcome, such as those promoting cell cycle arrest as opposed to RTC survival. Notably, the association with recurrence risk was stronger for the core RTC signature (HR=0.09) than the in vivo signature alone (HR=0.49)[7], and persisted following the removal of proliferation-associated genes (p=1.4e-16) (Fig. S1E). Thus, the core signature reflects features of cellular dormancy beyond known regulators of proliferation. Since >75% of patient tumors in this dataset were estrogen receptor positive (ER+), and more than two-thirds recurred at metastatic sites, our findings suggest that a conserved RTC signature generated from locally-derived, ER-negative RTCs in mice can identify mechanisms enabling RTC survival and recurrence for both distant recurrence and ER+ disease[7].

An in vivo ECM-focused loss-of-function screen identifies *B3galt6* as a regulator of RTC fitness

We reasoned that genes that were reversibly upregulated during dormancy could be involved in promoting RTC dormancy and/or survival. To filter this gene set and identify those specifically involved in RTC survival, we performed an in vivo CRISPR-Cas9-based loss-of-function screen targeting a clustered set of genes enriched for ECM-associated GO terms that displayed a reversible pattern of upregulation during in vitro dormancy (IVD) (Fig. 2A). We designed a custom sgRNA library targeting 95 candidate genes with 4–5 sgRNAs each (Table S1). An additional 53 sgRNAs included positive controls anticipated to be lethal to proliferating cells (*e.g.*, sg*Rpa3*, sg*Pcna*) and negative controls consisting of non-targeting sgRNAs and sgRNAs targeting inert sites (*e.g.*, sg*Rosa*) (Fig. 2B).

After transducing Her2-dependent-Cas9 cells with this GFP-labeled sgRNA library at low multiplicity of infection (MOI=0.3) and confirming that all sgRNAs were represented within GFP+ sorted cells (Fig. S2A), cells were orthotopically injected into mice. To assess sgRNA selection during disease progression we allowed Her2-driven PTs to form, then withdrew doxycycline to induce tumor regression to a non-palpable state (Fig. S2B) and harvested lesions at early (D7), mid (D14), and late (D28, D35) dormancy time points to assess sgRNA composition at each time point (Fig. 2B).

By sequencing the plasmid pool, sorted cells, injected cells, PTs and residual lesions (RLs), we confirmed that all sgRNAs were detectable at each time point (Fig. S2A). Calculating the Gini index to quantify the skewness of sgRNA distribution revealed that the plasmid pool, sorted cells, and injected cells displayed low Gini indices (median=0.21, 0.25, and 0.26, respectively), confirming homogeneous sgRNA distribution at these points. PTs displayed a slightly higher Gini index than pre-injection samples (median=0.29), suggesting only modest selection of sgRNAs in the presence of the strong oncogenic driver, Her2. In contrast, RLs displayed stepwise increases in Gini indices at D7, D14, D28, and D35 (median=0.38, 0.45, 0.58, 0.69, respectively), confirming that sgRNAs were progressively selected throughout dormancy (Fig. 2C).

To identify sgRNAs that underwent selection, we employed eight analytical methods and ranked sgRNAs identified by 2 or more methods. As anticipated, sg*Rpa3* and sg*Pcna* positive controls were depleted during PT formation (Fig. S2C), but not following Her2 downregulation, consistent with our prior findings that RTCs in this model are non-proliferative following Her2 downregulation[7] (Fig. S2D).

We reasoned that sgRNAs targeting genes that maintain cell cycle-arrest following Her2 downregulation would be enriched in RLs vs. PTs. Lending credence to this hypothesis, we found that sgRNAs targeting *Ddr1,* a known regulator of cell cycle-arrest[28], were enriched in RLs, which is consistent with a model in which *Ddr1* loss results in the expansion of RTCs (Fig. S2E). Conversely, we anticipated that sgRNAs targeting genes that maintain RTC survival would be depleted in RLs vs. PTs, as supported by the identification of the pro-survival gene, *Bcl2*, as one of the two top hits in this analysis (Fig. 2D).

The second top putative pro-survival hit, identified by all 8 methods in each of four pairwise comparisons of PT and RL time points, was *B3galt6* (Fig. 2D, S2F). Analyzing the distribution of sgRNAs targeting *B3galt6* (colored bars) vs. all sgRNAs (grey histogram) in each pairwise comparison confirmed that tumor cells harboring *B3galt6* sgRNAs were markedly depleted within 7 days following Her2 downregulation (PT vs. D7) and exhibited further depletion at later dormancy time points (D7 vs. D14, D28, D35) (Fig. 2E, S2F). This strong, persistent negative selection against cells harboring sg*B3galt6* suggests a potential role for B3GALT6 in dormancy.

B3GALT6 promotes RTC survival following therapy-associated dormancy

Beta-1,3-galactosyltransferase (B3GALT6) catalyzes the addition of the second Gal residue in the GlcA-Gal-Gal-Xyl-*O*- tetrasaccharide linker that is essential for the attachment of sulfated glycosaminoglycans (GAGs) to proteoglycan core proteins[29]. Accordingly, B3GALT6 enzymatic activity is required for proteoglycan assembly. Based on their linear chains of repeating disaccharides, GAGs are classified as: (i) heparan sulfate (HS); (ii) chondroitin sulfate (CS); or a CS derivative (iii) dermatan sulfate (DS) (Fig. 3A). Consequently, deletion of the *B3galt6* gene using CRISPR-Cas9 is predicted to ablate HS and CS/DS GAGs on cell surfaces[30].

To test this, we transduced Her2-dependent-Cas9 cells with sg*Rosa* or sg*B3galt6* vectors expressing GFP, performed ICE analysis[31], and confirmed that sg*B3galt6* efficiently generated indels predicted to result in loss-of-function mutations in >85% of tumor cells (Fig. S3A, B). Immunofluorescence demonstrated a marked reduction in HS levels in sg*B3galt6* cells vs. sg*Rosa* controls, providing further confirmation that these sgRNAs functionally reduce B3GALT6 activity (Fig. 3B, C).

To validate and extend findings from our CRISPR-Cas9 screen, we performed an IVD competition assay in which sg*Rosa*-GFP and sg*B3galt6*-GFP transduced Her2-dependent-Cas9 cells were plated at a 1:1 ratio with sg*Rosa*-mCherry cells on D-3. Genomic DNA was harvested from cells at D-2 or D0 (*baseline*), at D3, D7, D14, D21, and D28 following doxycycline withdrawal (*deinduction*), and at 48 hr after doxycycline re-addition to D28 deinduction cells (D28+, *reinduction*). Changes in %sg*B3galt6*-GFP:sg*Rosa*-mCherry cells relative to %sg*Rosa*-GFP:sg*Rosa*-mCherry controls were quantified by droplet digital PCR (ddPCR) (Fig. 3D). This revealed that sg*B3galt6* tumor cells were progressively depleted ($p<0.0001$) throughout dormancy beginning as early as 3 days following Her2 downregulation (Fig. 3E), indicating that tumor cell-autonomous B3GALT6 loss impairs RTC fitness.

To confirm that this dormancy-selective depletion of sg*B3galt6* cells occurs in vivo, as predicted by our CRISPR-Cas9 screen, we performed a competition assay analogous to that performed in vitro. sg*Rosa*-mCherry Her2-dependent-Cas9 cells admixed with an equal number of sg*Rosa*-GFP or sg*B3galt6*-GFP cells were injected into mice maintained on doxycycline, following which PTs (*baseline*) as well as D10 (*early*) and D28 (*late*) RLs were harvested (Fig. 3F).

While no differences in GFP intensity were visible in PTs across groups by whole-mount imaging, decreases in GFP intensity were evident for each of the sg*B3galt6* guides vs. sg*Rosa* controls in RL samples (Fig. 3F). ddPCR analysis confirmed this pronounced negative selection against sg*B3galt6* tumor cells (p<0.0001) in RLs and identified ongoing depletion from D10 to D28 (Fig. 3G). In contrast, sg*Rosa*-GFP cells exhibited no selection across time points. These data provide further evidence that B3GALT6 is required for maintaining RTC fitness in vivo.

To determine the cellular mechanisms by which B3GALT6 regulates RTC fitness in vivo, sg*B3galt6*_3-GFP or sg*Rosa*-GFP cells were injected into mice, followed by harvest of PTs and early (D4, D7) RLs following doxycycline withdrawal. As anticipated, a dramatic decrease in cell proliferation was observed in control lesions at D4 and D7 following doxycycline withdrawal as indicated by EdU incorporation (Fig. 3H, I) and Ki67 staining (Fig. S3C, D). No differences were observed in the %EdU+ or %Ki67+ RTCs between sg*B3galt6* and sg*Rosa*-derived PTs or RLs (Figs. 3H, I; S3C, D). These findings confirm that RTCs in both groups enter a quiescent state following Her2 downregulation, suggesting that depletion of sg*B3galt6* RTCs vs. control cells following Her2 downregulation is not attributable to differences in proliferation.

To determine if the decrease in sg*B3galt6* vs. *sgRosa* RTCs following Her2 downregulation resulted from differential cell survival, we quantified apoptosis using terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) (Fig. 3J, K) and cleaved caspase 3 (cc3) immunofluorescence (Fig. S3E, F). While no differences were apparent in PTs or D4 RLs between genotypes, a significant increase in %TUNEL+ (p=0.0009) and %cc3+ tumor cells (p=0.033) was evident at D7 in sg*B3galt6* vs. sg*Rosa* RL samples. These data indicate that the impaired fitness observed in sg*B3galt6* tumor cells following Her2 downregulation is attributable to higher levels of apoptosis in dormant sg*B3galt6* RTCs.

B3GALT6 promotes tumor recurrence following therapy-associated dormancy

In light of our observation that B3GALT6 promotes RTC fitness, we wished to determine whether B3GALT6 impacts the kinetics of spontaneous recurrence. Her2-dependent-Cas9 cells transduced with sg*B3galt6*-GFP or sg*Rosa*-GFP at a target MOI=5 (Fig. S4A, B) were injected into mice maintained on doxycycline to drive PT formation (Fig. 4A). sg*B3galt6* and sg*Rosa*-GFP groups displayed no differences in the rate of PT formation (p=0.26) (Fig. S4C).

Following doxycycline withdrawal to induce tumor regression to a non-palpable state, mice were monitored for recurrence. Strikingly, dormant RLs derived from either sg*B3galt6*_3 (p<0.0001; HR=0.13) or sg*B3galt6*_1 (p<0.0001; HR=0.11) tumor cells displayed dramatically delayed tumor recurrence vs. sg*Rosa* controls (Fig. 4B). Moreover, whereas all 20 recurrent tumors in the sg*Rosa* control group were strongly GFP+, only

1/20 sg*B3galt6*_3 and 3/20 sg*B3galt6*_1 recurrent tumors were GFP+ (Fig. 4C, S4D). This negative selection against sg*B3galt6* GFP+ cells in recurrent tumors was confirmed by ddPCR, which demonstrated >10-fold depletion of GFP+ cells in sg*B3galt6*_3 (p=0.0003) and sg*B3galt6*_1 (p=0.039) recurrences (Fig. 4D). The marked delay in spontaneous tumor recurrence in sg*B3galt6* cells, coupled with the observation that ~10% of untransduced GFP-negative cells efficiently and reproducibly outcompeted GFP+ sg*B3galt6* tumor cells during recurrent tumor formation, indicate that B3GALT6 is required for tumor recurrence.

To ensure that recurrences maintained B3GALT6 deletion, we transduced Her2-dependent-Cas9 cells with sg*Rosa*-GFP or sg*B3galt6*-GFP at a higher target MOI=10 (Fig. 4E), which resulted in a transduction efficiency of >95% (Fig. S4E, F). Under these conditions, sg*B3galt6*-GFP cells exhibited an even more dramatic delay in median recurrence-free survival (135d) vs. sg*Rosa*-GFP (39.5d) control cells (p<0.0001; HR=0.16) (Fig. 4F) than that observed for cells transduced at lower MOI (86–91d vs. 51d) (Fig. 4B). Furthermore, nearly all recurrent tumors arising from sg*B3galt6*-GFP (12/14) or sg*Rosa*-GFP (18/19) tumor cells were GFP+ (Fig. 4G, S4G) and ddPCR confirmed the presence of comparable numbers of GFP+ cells in sg*Rosa* and sg*B3galt6* recurrent tumors (Fig. 4H). As anticipated, immunofluorescence analysis demonstrated suppressed HS levels in sg*B3galt6* vs. sg*Rosa* recurrent tumors, confirming that loss of B3GALT6 function was maintained in sg*B3galt6* recurrences (Fig. S4H, I). These observations suggest that transducing tumor cells with sg*B3galt6* at a high MOI eliminated a bypass pathway in which untransduced, GFP-negative cells give rise to recurrence.

Analysis of EdU incorporation in recurrent tumors failed to identify differences in proliferation between sg*Rosa* and sg*B3galt6* MOI=10 recurrences (Fig. 4I, J). In contrast, TUNEL revealed increased apoptosis in sg*B3galt6* vs. sg*Rosa* recurrent tumors (p=0.018) (Fig. 4K, L). Consistent with these findings, a trending decrease in growth rate of ~2-fold was observed for sg*B3galt6* vs. sg*Rosa* recurrences (p=0.073) (Fig. 4M). These findings indicate that B3GALT6 is required for efficient tumor recurrence from RTCs post-therapy.

B3GALT6 promotes tumor cell survival and outgrowth in microenvironment-induced dormancy

Our data to this point identified B3GALT6 as a critical regulator of RTC survival and recurrence in the context of therapy-associated dormancy. We next asked whether B3GALT6 might also play a functional role in microenvironment-induced dormancy using the D2.OR-D2A1 paired cell line model. D2.OR-D2A1 cells share a common origin and grow comparably in two-dimensional (2D) culture, but manifest divergent growth properties (*i.e.*, D2.OR: dormant/indolent vs. D2A1: proliferative/aggressive) when cultured in 3D or at metastatic sites in vivo[32–34].

We first determined whether a dormancy-associated gene expression signature previously generated by comparing D2.OR to D2A1 cells in 3D[34] exhibited overlap with the core RTC signature derived from our therapy-associated model of dormancy (Fig. S1A, 1D). We observed significant overlap between these two signatures for both upregulated (hypergeometric p=1.68e-16) and downregulated (hypergeometric p=6.65e-03) genes (Fig. S5A). Furthermore, the D2.OR/D2A1 dormancy signature exhibited strong and progressive temporal enrichment within the gene expression changes that occurred in Her2-dependent tumor cells over the 28-day course following Her2 downregulation in vitro (Fig. 5A, p=1.88e-09), and was reversible following doxycycline readdition (D7+, D14+, D28+) (Fig. 5A). The D2.OR-derived dormancy signature was also enriched in a dormancy-specific manner in *MTB/TAN* and *MTB/TWNT* RTCs in vivo isolated 28 days after oncogene downregulation compared to either PTCs (p=7.4e-05, 1.7e-05, respectively) or recurrent tumor cells (p=6.2e-04, 1.1e-03, respectively) (Fig. S5B).

Next, we asked what GO terms were enriched within the D2.OR-derived dormancy signature. As observed for the core RTC signature derived from therapy-associated dormancy models, the D2.OR dormancy signature was enriched for multiple ECM, proteoglycan, and GAG-related GO terms (Fig. 5B). These data indicate that dormancy-associated gene expression signatures are strongly enriched for ECM-related pathways in general, and proteoglycan and GAG-related pathways in particular, irrespective of whether the dormant state was induced by the microenvironment or associated with therapy.

D2.OR cells persist as dormant solitary cells over the course of 12 days when grown on basement membrane extract (BME), whereas D2A1 cells maintained under these same conditions generate proliferative, spindle-shaped outgrowths[32]. Addition of Collagen I (Col I) to BME induces dormant D2.OR cells to resume proliferation and generate D2A1-like spindle-shaped colonies[32]. Using this system, we asked whether the dormant behavior of D2.OR cells in 3D culture is regulated by B3GALT6.

D2.OR and D2A1 cells transduced with a sh*B3galt6* hairpin exhibited >80% knockdown of *B3galt6* transcripts (Fig. S5C). D2.OR cells transduced with sh*B3galt6* or control sh*Scrambled* hairpins were overlaid on BME with or without Col I. While the numbers of viable sh*B3galt6* and sh*Scrambled* cells each decreased over the 12-day course, *B3galt6* knockdown resulted in progressive and marked decreases in cell number at D8 (p=0.022) and D12 (p=0.0004) compared to control cells (Fig. 5C, D).

As anticipated, sh*Scrambled* D2.OR cells plated on BME+Col I displayed increased numbers of viable cells over this same 12-day time course (Fig. 5E, F). In contrast, *B3galt6* knockdown in D2.OR cells yielded marked reductions in the numbers of viable tumor cells at D4, D8, and D12 compared to sh*Scrambled* controls (p<0.0001). Indeed, the Col I-induced increase in the number of viable D2.OR cells was entirely abrogated by *B3galt6* knockdown (Fig. 5E, F). Additionally, while sh*Scrambled* D2.OR cells formed spindle-like colonies when grown on BME+Col I, sh*B3galt6* cells failed to do so. Thus, B3GALT6 is required for both the viability of dormant D2.OR tumor cells and their outgrowth.

We performed analogous 3D assays using D2A1 cells. When grown on BME, a significant decrease in viable cell number was observed for sh*B3galt6* vs. sh*Scrambled* D2A1 cells at D12 (p=0.003), and sh*B3galt6* D2A1 cells failed to exhibit the characteristic spindle-like morphology adopted by sh*Scrambled* D2A1 cells that is associated with colony outgrowth (Fig. S5D, E). Notably, this same difference in phenotype was observed in sh*B3galt6* D2A1 cells (Fig. S5F) grown on BME+Col I. Moreover, a progressive decrease in viable cell numbers was observed for sh*B3galt6* D2A1 cells grown on BME+Col I at D4 (p=0.0014), D8 (p=0.0002), and D12 (p<0.0001) compared to control cells (Fig. S5G). These data indicate that, B3GALT6 is required for cell viability, as well as outgrowth, in proliferative D2A1 cells. In aggregate, these findings suggest that B3GALT6 is a critical regulator of dormant RTC survival, irrespective of whether dormancy is therapy-associated or induced by the microenvironment.

Heparan sulfate synthesis is upregulated during dormancy and associated with poor outcomes in patients with breast cancer

Because B3GALT6 catalyzes the synthesis of a linker that is common to, and required for, the production of both HS and CS/DS (Fig. 3A), we assessed the biosynthesis of these GAGs under dormancy conditions. First, we applied HS and CS/DS KEGG biosynthesis signatures to IVD gene expression data derived from Her2-dependent cells. This revealed that enzymes involved in HS synthesis are upregulated (Fig. S6A) – whereas those involved in CS/DS synthesis are downregulated (Fig. S6B) – during dormancy in vitro.

Addition of the hexosamine residue GlcNAc to the proteoglycan tetrasaccharide linker results in HS synthesis, whereas addition of GalNAc to this linker results in CS/DS synthesis[35]. Consistent with the KEGG biosynthetic signature analysis above, the enzymes that direct HS synthesis (*Extl2*, *Extl3*; highlighted in blue) were reversibly upregulated during dormancy (Fig. S6C); in contrast, the enzymes that direct CS/DS synthesis (*Csgalnact1*, *Csgalnact2*; highlighted in yellow) were reversibly downregulated during dormancy (Fig. S6D). Although not all biosynthetic enzymes were included in our CRISPR-Cas9 screen, we identified the HS polymerase *Ext2* as a hit, wherein sgRNAs targeting *Ext2* were selected against during dormancy, as were sgRNAs targeting *B3galt6* (Fig. 2D). In contrast, no negative selection was observed in our screen for sgRNAs targeting the CS polymerase *Chpf* (Fig. S2E). These associations suggest the intriguing possibility that HS is preferentially synthesized in dormant RTCs.

To quantify levels of HS and CS GAGs as a function of dormancy in vitro, we performed liquid chromatography-mass spectrometry (LC/MS) on Her2-dependent cell lysates isolated at baseline (D0, proliferative) or post-doxycycline withdrawal (D7, dormant). HS was readily detectable at D0 and its levels were significantly higher at D7 (p<0.0001) (Fig. 6A). In contrast, CS levels were >100-times lower than HS at baseline (D0) and were unchanged during dormancy (Fig. 6A).

We next evaluated GAG levels in vivo as a function of dormancy using HS or CS-specific antibodies. We first confirmed the specificity of these antibodies by demonstrating that pre-treatment of tissue sections with heparin lyase or chondroitinase eliminated the signal detected for HS or CS, respectively (Fig. S6E, G). Immunofluorescence staining revealed that Her2-dependent PTCs expressed HS (Fig. 6B, C; S6E, F), but not CS (Fig. 6D, E; S6G, H). Moreover, HS levels increased at D7 (Fig. 6B, C), whereas CS remained undetectable (Fig. 6D, E). Thus, both LC/MS and immunofluorescence data demonstrate an increase in HS abundance during dormancy.

Defects in differentiation resulting from impaired HS biosynthesis in myoblasts[36] and mouse embryonic stem cells[37] in vitro have been reported to be partially rescued by the addition of exogenous heparin, a highly sulfated variant of HS. Therefore, we performed an analogous experiment to determine whether heparin addition could rescue the impaired cell survival observed in B3GALT6-depleted cells, as would be predicted if HS promotes the survival of dormant tumor cells.

We harvested Her2-dependent sg*Rosa* and sg*B3galt6* tumor cells at baseline (D0, proliferative) and in cells treated with vehicle, low-dose heparin, or high-dose heparin for 7 days post-doxycycline withdrawal (D7, dormant). As above, sg*B3galt6* cells exhibited decreased survival at D7 vs. sg*Rosa* (Fig. 6F). Consistent with reports that heparin supplementation in HS-replete cells can dampen signaling by competing with endogenous HS[38], we observed that both low- and high-dose heparin treatment in sg*Rosa* cells modestly decreased viable RTC number compared to vehicle-treated controls (Fig. 6F). In contrast, exogenous addition of high-dose heparin to sg*B3galt6* RTCs markedly increased the number of viable tumor cells at D7 vs. vehicle-treated controls ($p<0.0001$) (Fig. 6F). These data indicate that heparin supplementation is sufficient to rescue the survival defect induced by loss of B3GALT6 in RTCs.

To this point, our functional data in preclinical models suggested a selective and essential role for HS, but not CS, in dormant tumor cell survival. Therefore, we applied the KEGG HS and CS/DS biosynthetic signatures to gene expression data from primary breast cancers in ~4400 patients with known recurrence outcomes[20]. Consistent with a role for HS in enhancing tumor cell fitness during dormancy and recurrence, the expression of HS biosynthetic enzymes in early-stage primary breast cancers was strongly associated with poorer recurrence-free survival (overall p value=2.2e-04) (Fig. 6G). In contrast, no association was observed between recurrence-free survival and expression of CS/DS biosynthetic enzymes (Fig. 6H). These data provide further evidence specifically implicating HS biosynthesis in promoting breast cancer recurrence in patients.

Heparan sulfate 6-*O*-sulfation is selectively upregulated during dormancy and potentiates FGF1 signaling

Having identified a key role for B3GALT6-mediated HS biosynthesis in promoting RTC survival, we asked if HS undergoes dynamic modifications during dormancy, as suggested by the reversible upregulation of gene expression for sulfotransferases (*e.g.*, *Ndst2*, *Hs2st1*, *Hs6st1*) within the HS signature (Fig. S6C). Accordingly, we examined the sites and extent of heparan sulfation since this modification is a key determinant of the ligand binding properties of GAG side chains[39,40].

To accomplish this, we isolated cell lysates from IVD samples at D0 (*baseline*), D7 and D28 (*deinduction*), and D28+ (*reinduction*), and performed glycan reductive isotope labeling followed by LC/MS (GRIL-LC/MS)[41] (Fig. S7A). This revealed a significant increase in the average sulfation level per HS-derived disaccharide at D7 and D28 vs. D0 baseline ($p<0.0001$), which remained elevated after 72 hr of Her2 reinduction (Fig. S7B). This increase in sulfation indicates that, in addition to the increased overall levels of HS GAGs, the extent to which these GAGs are sulfated is also dynamically upregulated in dormant RTCs.

Next, we wished to determine whether dormancy-selective patterns of heparan sulfation were present. Of the three (out of four total) sulfation sites on HS-derived disaccharides that we interrogated (Fig. 7A–D; S7C), 6-*O*-sulfation exhibited the most dramatic fold-increase in abundance during dormancy (Fig. 7D). We also examined the relative abundance of specific disaccharide motifs[42] as a function of dormancy (Fig. S7D, E). Only D2S6 and D0A6 disaccharides displayed upregulation during dormancy in a manner that was reversible following exit of cells from the dormant state (Fig. S7F, G). This suggests that 6-*O*-sulfation on glucosamine tends to co-occur in non-sulfated uronic acid-*N*-acetylated glucosamine (D0A6) or 2-*O*-sulfated uronic acid-*N*-sulfated glucosamine (D2S6) disaccharides.

Heparan 6-*O*-sulfation promotes FGF signaling by increasing the binding affinity of some FGF ligands for HS in the ternary HS-FGF-FGFR complex that is required for FGF signaling[43–45] (Fig. S7H). To identify which FGFs are expressed in dormant tumor cells in an abundant, but reversible, manner we interrogated the expression of the 15 paracrine FGFs that require heparin/HS as a co-factor for signaling[46]. This revealed that FGF1 is dramatically upregulated (>200-fold) during dormancy as early as 3 days following doxycycline withdrawal, and that its expression is potently suppressed within 48 hr of doxycycline re-addition (D7+, D14+, D28+) (Fig. 7E). In contrast, while FGF2 and FGF7 demonstrated dormancy-associated expression, they did not meet our criteria of being both abundantly expressed and reversibly upregulated (Fig. S7I).

To test whether endogenous FGF1 regulates RTC survival, we determined the number of dormant RTCs that survived Her2 downregulation following depletion of endogenous FGF1. After confirming successful editing by *Fgf1* sgRNAs (Fig. S7J), we performed an IVD assay using Her2-dependent-Cas9 cells transduced with sg*Fgf1* or sg*Rosa*. Deletion of *Fgf1* further accentuated the decrease in viable RTC numbers that occurs at D7 (*deinduction*) vs. D0 (*baseline*) when compared to sg*Rosa* controls (Fig. S7K). This suggests that tumor cell-autonomous FGF1 signaling promotes dormant RTC survival.

To determine whether FGF acts via heparan 6-*O*-sulfation to mediate B3GALT6-dependent RTC survival, we examined the expression of the three sulfotransferases, *Hs6st1–3,* that catalyze 6-*O*-sulfation (Fig. 7F; S7H, L), particularly those on the D2S6 and D0A6 6-*O*-sulfated disaccharide motifs[40,43] identified above. Of these, only *Hs6st1* was abundantly expressed and reversibly upregulated during dormancy in Her2-dependent cells (Fig. 7F). Notably, *Hs6st1* was identified as a putative pro-survival hit in our CRISPR-Cas9 screen, and the magnitude of its effects appeared to increase at later dormancy time points (Fig. 2D, 7G).

Next, we asked if impaired HS biosynthesis or 6-*O*-sulfation in RTCs results in an attenuated viability response to FGF1, as well as whether this could be rescued by the addition of heparin, which is natively sulfated including at the 6-*O* site. After confirming successful editing by *Hs6st1* sgRNAs (Fig. S7M), we treated Her2-dependent tumor cells with FGF1, heparin, or a combination of FGF1 + heparin for 7 days post-doxycycline withdrawal. We found that sg*B3galt6* and sg*Hs6st1* RTCs (D7, dormant) display reduced viability in response to exogenous FGF1 stimulation compared to sg*Rosa* RTCs, and that heparin could rescue this blunted response (Fig. 7H).

To extend these findings to a microenvironment-induced model of dormancy, we plated sh*B3galt6* or sh*Scrambled*-transduced D2.OR cells in 3D in the presence of vehicle, FGF1, heparin or a combination of FGF1 + heparin for 10 days. Loss of B3GALT6 resulted in a pronounced defect in the viability of dormant D2.OR cells grown on BME in the presence of FGF1 (Fig. 7I, K). Notably, addition of exogenous heparin rescued this decrease in viable cell number observed in dormant D2.OR sh*B3galt6* cells vs. sh*Scrambled* (p<0.0001) (Fig. 7I, K). Additionally, when D2.OR cells were plated under outgrowth conditions (BME+Col I), heparin potently rescued the decrease in cell numbers observed in FGF1-treated dormant D2.OR sh*B3galt6* vs. sh*Scrambled* cells (p<0.0001) (Fig. 7J, K). Furthermore, the combination of FGF1+heparin rescued the rounded morphology associated with sh*B3galt6*, suggesting that FGF1 and heparin may additionally stimulate D2.OR cell outgrowth (Fig. 7J). Thus, data from these two independent dormancy systems are consistent with a model wherein heparan 6-*O*-sulfation potentiates FGF signaling to promote the survival of dormant RTCs.

FGFR2 is selectively upregulated during dormancy and promotes ERK1/2 signaling

To determine the mechanism by which FGF1 signaling promotes dormant RTC survival, we evaluated the expression of the four FGFRs, *Fgfr1–4*, that participate in FGF signaling. Of these, *Fgfr1* and *Fgfr2* are abundantly expressed in Her2-dependent cells and are each reversibly upregulated during dormancy (Fig. 8A, S8A); *Fgfr2* was upregulated ~4-fold during dormancy, whereas *Fgfr1* was upregulated ~1.5-fold. Although FGF1 can signal through both FGFR1 and FGFR2, *Fgfr2* was identified as a putative pro-survival hit in our CRISPR-Cas9 screen (Fig. 2D, 8B), whereas sgRNAs targeting *Fgfr1* displayed no selection (Fig. S2E).

To determine whether these findings also applied to other cell models, we treated ER+ MCF7 and Her2-amplified BT474-M1 human breast cancer cell lines, with Fulvestrant, and Lapatinib, respectively. We identified a graded temporal increase in *Fgfr2* expression following targeted therapy with either anti-estrogen or anti-Her2 pharmacological agents (Fig. 8C), suggesting that FGFR2 signaling may constitute a generalized survival response in RTCs. Furthermore, we found that FGF1 stimulation increased cell viability in Fulvestrant-treated MCF7 and Lapatinib-treated BT474-M1 tumor cells (Fig. 8D). Accordingly, we hypothesized that FGFR2 may act together with 6-*O*-sulfated HS to promote FGF1 signaling and RTC survival, and prioritized studies of FGFR2 as a candidate for maintaining dormant RTC survival.

After confirming successful editing by *Fgfr2* sgRNAs (Fig. S8B), we performed an IVD assay using Her2-dependent-Cas9 cells transduced with sg*B3galt6*, sg*Hs6st1*, sg*Fgfr2*, or sg*Rosa*. We reasoned that if increased HS synthesis and 6-*O*-sulfation by dormant RTCs results in enhanced endogenous FGF signaling, then impairing HS synthesis (sg*B3galt6*), or 6-*O*-sulfation (sg*Hs6st1*), should attenuate FGF signaling (Fig. S8C). Furthermore, we predicted that if enhanced FGF signaling in dormant cells is mediated by FGFR2, depleting *Fgfr2* should also attenuate FGF signaling during dormancy.

To test this hypothesis, we assessed levels of activated ERK1/2[47] in each of the above genetic contexts, using two sgRNAs each for sg*B3galt6*, sg*Hs6st1*, and sg*Fgfr2*. As anticipated, pERK1/2 levels normalized to total ERK1/2 were markedly higher (>10-fold) in Her2-dependent proliferating cells at D0 (*baseline*), under which conditions the doxycycline-dependent bicistronic Her2-IRES-luciferase transgene is abundantly expressed (Fig. S8D) compared to D4 or D7 dormancy time points (Fig. 8E). Notably, pERK1/2:ERK1/2 levels in the presence of Her2 expression (D0) were largely unaffected by deletion of *B3galt6*, *Hs6st1*, or *Fgfr2* (Fig. 8E). In contrast, pERK1/2:ERK1/2 levels were significantly diminished in sg*B3galt6*, sg*Hs6st1*, and sg*Fgfr2* cells compared to *sgRosa* controls following Her2 downregulation (*deinduction*) resulting from doxycycline withdrawal (Fig. 8F). This impairment in pERK1/2:ERK1/2 levels in dormant RTCs caused by deletion of components of the FGF signaling pathway, which are upregulated in RTCs in a dormancy-specific manner, is consistent with a model in which endogenous FGF signaling is active during dormancy and requires B3GALT6, HS6ST1, and FGFR2 for its maintenance.

Upregulation of heparan sulfate 6-*O*-sulfation during dormancy promotes RTC survival and recurrence in part by promoting FGFR2 signaling

We wished to determine whether heparan 6-*O*-sulfation and FGF signaling are required for the viability of dormant RTCs, as suggested by our CRISPR screen results (Fig. 2D, 8B). We assessed viable RTC numbers for sg*Rosa*, sg*B3galt6*, sg*Hs6st1*, and sg*Fgfr2*-transduced Her2-dependent-Cas9 cells at D14 (*deinduction*) following doxycycline withdrawal. Consistent with our hypothesis, and concordant with their effects on FGF signaling, tumor cells deleted for each of these genes exhibited significantly impaired survival during dormancy compared to sg*Rosa* control cells (Fig. S8E). This suggests that heparan 6-*O*-sulfation potentiates endogenous FGF signaling mediated by FGFR2 to promote the survival of dormant RTCs in vitro.

To extend these findings in vivo, we injected GFP-labeled sg*Rosa*, sg*Hs6st1*, and sg*Fgfr2* Her2-dependent-Cas9 cells into mice. Her2-driven PTs (*baseline*) and D7 RLs following doxycycline withdrawal were harvested for TUNEL staining. This revealed a significant increase in the percentage of TUNEL+ sg*Hs6st1* (p=0.003) and sg*Fgfr2* (p=0.04) tumor cells in D7 RLs vs. sg*Rosa* controls (Fig. 8G, H). In contrast, we found no significant differences in the percentage of TUNEL+ tumor cells in sg*Hs6st1,* sg*Fgfr2,* and *sgRosa* PT samples (Fig. 8G, H). These data indicate that HS6ST1 and FGFR2 are selectively required for tumor cell survival in a dormancy-specific manner, which is analogous to our findings for B3GALT6.

To probe their respective pathway relationships, we asked whether HS 6-*O*-sulfation and FGFR2 act in concert to promote FGF1 signaling and dormant RTC survival. To accomplish this, we performed a competition assay by injecting sg*Rosa*-mCherry Her2-dependent-Cas9 cells admixed with an equal number of singly transduced (a) sg*Rosa*-GFP, (b) sg*B3galt6*-GFP, (c) sg*Hs6st1*-GFP, or (d) sg*Fgfr2*-GFP cells, or doubly transduced (e) sg*Rosa*-GFP+sg*B3galt6*-GFP, (f) sg*B3galt6*-GFP+sg*Hs6st1*-GFP, (g) sg*B3galt6*-GF+sg*Fgfr2*-GFP, or (h) sg*Hs6st1*-GFP+sg*Fgfr2*-GFP tumor cells, after confirming high editing efficiency

(Fig. S8F). These combinations of tumor cells were injected into mice, following which PTs (*baseline*) as well as D10 RLs (*early*) were harvested.

We hypothesized that the absence of HS biosynthesis in sg*B3galt6* RTCs would result in the strongest RTC fitness phenotype because of its broader impact on multiple growth factor-dependent pathways, including but not limited to FGF signaling. In the specific context of FGF-dependent RTC viability, we anticipated that FGFR2 deletion would result in a more potent phenotype than HS6ST1 deletion, given that FGFR2 is the direct mediator of FGF signaling whereas HS6ST1 modulates FGF avidity for FGFR2 binding.

In accordance with our hypothesis, we observed the greatest depletion of RTCs between PT and D10 RL for sg*B3galt6* tumor cells, followed by sg*Fgfr2*, and then sg*Hs6st1* (Fig. 8I). Furthermore, analysis of the phenotypes of double knockouts revealed that *B3galt6* was epistatic to both *Hs6st1* and *Fgfr2* (Fig. 8I, S8G). That is, combined deletion of *Hs6st1* and *B3galt6,* or of *Fgfr2* and *B3galt6,* exhibited effects that were no greater than the effect of deleting *B3galt6* alone. In an analogous manner, analysis of double knockouts revealed that *Fgfr2* was epistatic to *Hs6st1*, insofar as the effect of combined deletion of *Hs6st1* and *Fgfr2* was no greater than the effect of deleting *Fgfr2* alone (Fig. 8I, S8G). These data provide compelling in vivo evidence that B3GALT6, HS6ST1, and FGFR2 act within the same pathway, and that FGFR2 constitutes one of the mechanisms downstream of B3GALT6 that maintains RTC survival.

To determine the impact of HS6ST1 and FGFR2-mediated RTC pro-survival effects on tumor recurrence in vivo, we performed recurrence-free survival assays in mice bearing RLs derived from sg*Hs6st1* or sg*Fgfr2* Her2-dependent tumor cells. Analogous to our findings for sg*B3galt6*, mice bearing RLs containing sg*Hs6st1* or *Fgfr2* tumor cells exhibited delayed recurrence-free survival compared to sg*Rosa* controls (p=0.02, HR=0.40; and p=0.006, HR=0.35, respectively) (Fig. 8J). In aggregate, our studies suggest a model in which HS6ST1 is reversibly upregulated during dormancy, which results in the selective upregulation of heparan sulfate 6-*O*-sulfation in dormant RTCs and potentiation of FGF1 signaling via FGFR2, which promotes dormant RTC survival and tumor recurrence.

B3GALT6 pathway alterations in patients with breast cancer are associated with poorer clinical outcomes

Supporting the clinical relevance of alterations in the B3GALT6/HS6ST1/FGF1/FGFR2 pathway to human breast cancer progression, we found that amplification of at least one member of this set of genes (*B3GALT6* [All: n=5; HR$^+$ n=1], *HS6ST1* [All: n=1; HR$^+$ n=0], *FGF1* [All: n=3; HR$^+$ n=3)], or *FGFR2* [All: n=13; HR$^+$ n=8]) was associated with significantly reduced recurrence-free survival across all patients with breast cancer (p=0.02, HR=3.0), as well as in patients with hormone receptor-positive (HR+)/HER2-negative tumors (p=0.004, HR=4.7), which is the predominant breast cancer subtype (Fig. 8K). In particular, *FGFR2* copy number gain alone was significantly associated with reduced recurrence-free survival in patients with HR+/HER2-negative breast cancers (p=0.004, HR=1.7), and across all patients (p=0.055, HR=1.3) (Fig. 8L).

Consistent with the above associations, we found that FGF pathway activity was significantly enriched in residual tumors in four[21–23,25,26] of six patient datasets following neoadjuvant endocrine therapy or chemotherapy, and was co-enriched with a HS biosynthesis signature in two[21–23] of these neoadjuvant datasets. Moreover, analogous to our findings using a HS biosynthesis signature (Fig. 6G), we found that FGF pathway activity in early-stage primary breast cancers was significantly associated with poorer recurrence-free survival (p=0.021) (Fig. S8H).

Lastly, while we have not identified gene expression datasets comparing disseminated tumor cells (DTCs) to proliferating tumor cells in the same patients, we did evaluate a paired data set comparing gene expression levels in circulating tumor cells (CTCs) and metastatic breast cancers[48]. CTCs represent a population of tumor cells that can give rise to DTCs and, like DTCs, they are exposed to significant biological stresses encountered at foreign sites. Strikingly, analysis of this patient data set revealed that *B3GALT6*, *HS6ST1*, *FGF1*, and *FGFR2* are each highly and significantly upregulated in CTCs compared to their metastatic counterparts ~20-fold (p=0.004), 25-fold (p=0.005), 71-fold (p=4.58e-05), and 19-fold (p=0.013), respectively (Fig. 8M). These data further implicate alterations in a B3GALT6/HS6ST1/FGF1/FGFR2 pathway in human breast cancer progression.

Discussion

RTCs that persist following primary tumor therapy constitute the precursors of treatment-refractory recurrent disease that determines patient mortality. As some post-adjuvant clinical trials to identify and target RTCs in patients with early-stage breast cancer are currently underway[49,50], there is a pressing need to identify unique vulnerabilities of dormant RTCs that could be leveraged in such studies.

Here, we found that dormant RTCs create – in a cell-autonomous manner – an extracellular environment that is conducive to their own survival. In doing so, we identify a dependency of dormant RTC fitness that could be exploited to prevent tumor recurrence. Specifically, we determined that dormant RTCs selectively upregulate the B3GALT6-mediated synthesis of HS proteoglycans in both therapy-associated and microenvironment-induced models of dormancy. Mechanistically, RTCs reversibly upregulate B3GALT6-mediated HS synthesis, HS6ST1-mediated heparan 6-*O*-sulfation, as well as *Fgf1* and *Fgfr2* expression, in a dormancy-specific manner. These orchestrated effects result in enhanced FGF1-FGFR2 signaling, enhanced dormant RTC survival, and accelerated tumor recurrence.

Underscoring the clinical relevance of these data, we found that amplification of at least one member of the B3GALT6/HS6ST1/FGF1/FGFR2 axis was associated with significantly reduced recurrence-free survival, as was expression of a HS biosynthesis signature or a FGF pathway activity signature in patients with early-stage breast cancers. The concordance of these findings in patients with functional studies in mice further implicates the B3GALT6/HS6ST1/FGF1/FGFR2 axis as a potential target for preventing recurrence from dormant RTCs.

Our findings identify B3GALT6 as a potent regulator of dormant RTC survival and recurrence. To date, a role for B3GALT6 in cancer has not been reported. Because B3GALT6 is involved in the synthesis of the tetrasaccharide required for covalent linkage of GAG chains to the protein core[29], genetic deletion of *B3galt6* permits investigation of the role of GAGs in tumor progression. Prior studies of B3GALT6 in vertebrate models have been limited to genetic loss in a zebrafish model[30] generated for study of two pathogenic conditions observed in patients with biallelic loss of *B3GALT6*: spondylodysplastic Ehlers-Danlos syndrome (spEDS) and spondyloepimetaphyseal dysplasia with joint laxity, type 1 (SEMD-JL1)[51,52]. While zebrafish with *b3galt6* loss-of-function recapitulate the connective tissue defects observed in these patients, the extent to which the pathways underlying B3GALT6-mediated functions in connective tissues and dormant RTCs are shared – if at all – remains to be determined.

Heparan 6-*O* sulfation is required for formation of the ternary HS-FGF-FGFR complex and, therefore, FGF pathway activity[43,44]. In particular, 6-*O*-sulfation is a strong determinant of FGF1 binding[45]. Although FGF signaling is classically thought to play a mitogenic function in tumor cells[9,53], recent data suggest a context-dependent role for FGF signaling in promoting breast cancer therapy resistance and growth arrest. FGF2 derived from osteogenic cells in the bone marrow can suppress the expression of ER in DTCs via FGFR1, rendering them resistant to endocrine therapy[54]. Furthermore, exogenously-derived FGF2 induces cell cycle arrest in ER+ breast cancer cell lines[55] and can induce the expression of the pro-dormancy transcription factor ZFP281 in early disseminated cancer cells[56].

In contrast to these examples in which FGF2 promotes dormancy predominantly via FGFR1, our data suggest that dormant RTC survival is mediated by FGF1 in a HS-FGFR2-dependent manner. Consistent with our CRISPR-Cas9 screen identifying *Fgfr2*, but not *Fgfr1*, as a vulnerability in dormant RTCs, we found that FGFR2 deletion induced RTC apoptosis and significantly delayed tumor recurrence. These data suggest the intriguing possibility that a highly selective FGFR2 pharmacological inhibitor[57], currently in clinical trials for intrahepatic cholangiocarcinoma (NCT04526106[58]), could be leveraged in the adjuvant/post-adjuvant setting to exploit the dependency of dormant RTCs on FGF signaling for their survival and recurrence.

STAR Methods

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Lewis A. Chodosh, M.D., Ph.D. (chodosh@pennmedicine.upenn.edu)

Materials availability—All unique/stable reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- The RNA-sequencing dataset reported in this article is deposited in NCBI's gene expression omnibus (GEO) and is publicly available as of the date of publication. The accession number is listed in the Key Resources Table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All animal studies were approved by the University of Pennsylvania Institutional Animal Care and Use Committee (IACUC). Female *nu/nu* mice (NCRNU-F, Taconic) at 5–6 weeks of age were used as transplantation donors in this study, ear-tagged, and randomly assigned to experimental groups. Mice were housed in a barrier facility with 12-hr light/12-hr dark cycles, with access to food and water *ad libitum*.

Primary Her2-dependent tumor cells were derived and cultured from female *MTB/TAN* FVB transgenic mice as previously described[59] and confirmed to be mycoplasma-negative. *MTB/TAN*-derived primary tumor cells were cultured in DMEM (Corning, Cat. # 10-017-CV) containing 10% super calf serum (GeminiBio, Cat. # 100–510), 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122), 1% Glutamine (Thermo Fisher Scientific, Cat. # 25030081), 2mg/ml Doxycycline (RPI, Cat. # D43020–250.0), 5mg/ml Prolactin (NHPP, Cat. # NIDDK-oPRL-21), 5mg/ml Insulin (GeminiBio, Cat. # 700–112P), 10ug/ml EGF (Millipore, Cat. # E4127), 1mg/ml Hydrocortisone (Sigma, Cat. # H0396), 1mM Progesterone (Sigma, Cat. # P7556).

D2.OR and D2A1 cells were derived from female Balb/c mice and were a gift from Dr. Mikala Egeblad at Cold Spring Harbor Laboratories. These cells were maintained in in DMEM (Corning, Cat. # 10-017-CV) containing 10% fetal bovine serum (GeminiBio, Cat. # 100–106), 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122).

MCF7 cells (ER+, HER2-negative) were derived from a female patient with breast cancer and obtained as part of the NCI authenticated panel NCI ICBP43. MCF7 cells (ER+, HER2-negative) (ATCC, Cat. # HTB-22) were cultured in EMEM (ATCC, Cat. # 30–2003) containing 10% fetal bovine serum (GeminiBio, Cat. # 100–106), 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122), 0.01mg/ml Insulin (GeminiBio, Cat. # 700–112P).

BT474-M1 cells (ER+, HER2+) is a lung metastatic subclone of the parental BT-474 cells derived from a female patient with breast cancer and were a gift from Dr. Mien-Chie Hung at MD Anderson Cancer Center. Cells were cultured in DMEM/F12 (Hyclone, Cat. # SH30023.01) containing 10% fetal bovine serum (GeminiBio, Cat. # 100–106), 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122).

No STR testing was performed for authentication of these cell lines.

METHOD DETAILS

In vitro assays—*MTB/TAN*-derived primary tumor cells were cultured in DMEM (Corning, Cat. # 10-017-CV) containing 10% super calf serum (GeminiBio, Cat. # 100–510), 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122), 1% Glutamine (Thermo Fisher Scientific, Cat. # 25030081), 2mg/ml Doxycycline (RPI, Cat. # D43020–250.0), 5mg/ml Prolactin (NHPP, Cat. # NIDDK-oPRL-21), 5mg/ml Insulin (GeminiBio, Cat. # 700–112P), 10ug/ml EGF (Millipore, Cat. # E4127), 1mg/ml Hydrocortisone (Sigma, Cat. # H0396), 1mM Progesterone (Sigma, Cat. # P7556).

For in vitro dormancy experiments, cells were plated on D-3 in medium as above, transitioned to medium as above containing 1% super calf serum on D-2, and transitioned to medium containing 1% super calf serum and no Doxycycline on D0. Plates were harvested at dormancy deinduction time points as desired, while medium containing 1% super calf serum and no Doxycycline was replaced on remaining plates weekly. Reinduction plates were harvested at time points as desired following 48h or 72h treatment with medium containing 1% super calf serum and doxycycline as indicated in the figures. For FGF1 and heparin stimulation experiments, cells were treated daily starting at D0 with vehicle (PBS), mFGF1 (25ng/ml) (R&D Systems, Cat. # AF4686), heparin (25µg/ml) (Fisher Scientific, Cat. # AAA1619803), or mFGF1 + heparin and harvested on D7.

D2.OR and D2A1 cells were a gift from Dr. Mikala Egeblad at Cold Spring Harbor Laboratories. These cells were maintained in in DMEM (Corning, Cat. # 10-017-CV) containing 10% fetal bovine serum (GeminiBio, Cat. # 100–106), 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122).

MCF7 cells (ER+, HER2-negative) (ATCC, Cat. # HTB-22) were cultured in EMEM (ATCC, Cat. # 30–2003) containing 10% fetal bovine serum (GeminiBio, Cat. # 100–106), 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122), 0.01mg/ml Insulin (GeminiBio, Cat. # 700–112P). Following 24h of plating, the medium was replaced with serum-starved medium 1% fetal bovine serum. Following 48h of serum starvation, cells were treated daily with vehicle (DMSO) or 100nM Fulvestrant (Selleck Chemicals, Cat. # S1191) for 72h and cells were harvested for qRT-PCR. For the FGF1 stimulation experiment, 100nM Fulvestrant-treated cells were treated daily with vehicle (PBS) or 25ng/ml hFGF1 (R&D Systems, Cat. # 232-FA) for 72h prior to harvest.

BT474-M1 cells (ER+, HER2+) were a gift from Dr. Mien-Chie Hung at MD Anderson Cancer Center. Cells were cultured in DMEM/F12 (Hyclone, Cat. # SH30023.01) containing 10% fetal bovine serum (GeminiBio, Cat. # 100–106), 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122). Following 24h of plating, the medium was replaced with serum-starved medium 1% fetal bovine serum. Following 48h of serum starvation, cells were treated daily with vehicle (DMSO) or 400nM Lapatinib (Selleck Chemicals, Cat. # S2111) for 72h and cells were harvested for qRT-PCR. For the FGF1 stimulation experiment, 400nM Lapatinib-treated cells were treated daily with vehicle (PBS) or 25ng/ml hFGF1 (R&D Systems, Cat. # 232-FA) for 72h prior to harvest.

Cell viability assays (2D and 3D) were performed as per manufacturer's instructions using the Cell Titer 96 Non-Radioactive Cell Proliferation Assay (Promega, Cat. # G4000) in 96-well plates.

3D assays—D2.OR and D2A1 cells were plated on top of a matrix of basement membrane extract (Cultrex 3D Basement Membrane Extract, Reduced Growth Factor, Trevigen Cat. # 3445-005-01) or a 1:1 ratio of basement membrane extract and neutralized type I collagen (Cultrex 3D Culture Matrix Rat Collagen I, Trevigen Cat. # 3447-020-01). For experiments in 96-well plates (Corning, Cat. # 353219), 50µl of the matrix was added per well and were solidified for 1h at 37°C. D2.OR and D2A1 cells were resuspended at 20,000 cells/ml in DMEM low glucose, low pyruvate medium (Thermo Fisher Scientific, Cat. # 11885092) containing 2% fetal bovine serum (GeminiBio, Cat. # 100–106), 2% basement membrane extract, 1% Penicillin/Streptomycin (Thermo Fisher Scientific, Cat. # 15-140-122), and 100µl per well was plated on top of the solidified matrices. Plates were harvested at desired time points and medium was replaced on remaining plates every 4 days. For FGF1 and heparin stimulation experiments, cells were treated daily starting at D0 with vehicle (PBS), mFGF1 (25ng/ml) (R&D Systems, Cat. # AF4686), heparin (25µg/ml) (Fisher Scientific, Cat. # AAA1619803), or mFGF1 + heparin and harvested on D10. Cell viability in the harvested plates was assayed using the CellTiter Non-Radioactive Cell Proliferation Assay (Promega, Cat. # G4000) as per manufacturer's instructions.

RNA sequencing—*MTB/TAN* cells were cultured in vitro under dormancy conditions as described. Samples were harvested at baseline (D0), deinduction (8 h, D1, D3, D7, D14, D28), and reinduction (D7+, D14+, and D28+) timepoints. RNA isolation was performed using the RNeasy Mini kit (Qiagen, Cat. # 74106) and sequencing libraries were prepared using the TruSeq Stranded mRNA for NeoPrep kit (Illumina, Cat. # NP-202-1001). Sequencing was performed using 75-bp paired-end NextSeq (Illumina, Cat. # 20024907).

The quality of raw reads was assessed using FASTQC. Sequenced reads were mapped to the mm10 *Mus musculus* reference genome using Spliced Transcripts Alignment to a Reference (STAR). Gene-level read counts were determined using featureCounts. Principal component analysis (PCA) was performed using the top 1000 most variable genes assessed by per-gene standard deviations and was used to exclude distinct outliers from the downstream analysis. Read counts across samples were normalized and differentially expressed genes were identified using DESeq2.

Animal experiments—All animal studies were approved by the University of Pennsylvania Institutional Animal Care and Use Committee (IACUC). In vivo competition assays and recurrence-free survival assays were performed as previously described. Briefly, 1e06 cells (1:1 mixture of GFP$^+$ and mCherry$^+$ cells for competition assays or unmixed cells for recurrence-free survival assays and the CRISPR-Cas9 screen) were transplanted into the inguinal #4 mammary fat pads of *nu/nu* female mice (NCRNU-F, Taconic) and administered 2mg/ml doxycycline/5% sucrose via drinking water. Mice were de-induced by switching to regular drinking water once mammary tumors reached target size (5×5mm for competition assays; 3×3mm for recurrence-free survival assays and the CRISPR-Cas9 screen). Mice were palpated thrice a week and time-to-recurrence, as measured by when non-palpable, dormant lesions first became palpable, was assessed by Kaplan-Meier analysis. At the time of harvest, mice were administered 50mg/kg EdU (i.p.) 2h prior to euthanasia.

Plasmids and lentiviral production—LentiV_Cas9_puro (Addgene, Cat. # 108100) was used to generate Cas9 expressing *MTB/TAN*-derived primary tumor cells. LRG2.1 (Addgene, Cat. # 108098), LRG (Addgene, Cat. # 65656), or LRmCherry2.1 (Addgene, Cat. # 108099) vector backbones were utilized for cloning sgRNAs for CRISPR-Cas9 studies[60]. For each sgRNA, sense and antisense oligos were phosphorylated and annealed and then ligated into BsmB1-digested vector. Ligated vectors were transformed into the chemically competent Stbl3 bacteria (Thermo Fisher Scientific, Cat. # C737303). Successfully transformed bacterial clones were picked from Ampicillin selective plates and isolated DNA was sequenced using a U6 primer to confirm sgRNA incorporation.

Lentiviruses were generated in HEK293T cells using the TransIT-293 transfection reagent (Mirus, Cat. # MIR2700) to introduce packaging plasmids pMD2.G (Addgene, Cat. # 12259; 3µg) and psPAX2 (Addgene, Cat. # 12260; 6µg), and 9µg of the desired backbone containing the sgRNA of interest. sgRNA lentiviruses were titered by serial dilution in *MTB/TAN*-derived primary tumor cells using the fluorophore associated with the vector backbones as a readout by on the Attune NxT flow cytometry (Thermo Fisher).

PCR and ICE analysis—To validate sgRNA editing efficiency, genomic DNA was isolated from Her2-dependent Cas9 unedited and edited cells one week after transduction with sgRNAs (Qiagen, QIAamp DNA mini kit, Cat. # 51304). PCR amplification was performed (98°C 5min → (98°C 5s, T_m 5s, 72°C 15s) ×30 → 72°C 3min → 4°C) using the following primer sequences and annealing temperatures (T_m):

sg*Rosa*: F-GCGGGAGAAATGGATATGAA; R-GCACTTGCTCTCCCAAAGTC; T_m = 60°C

sg*B3galt6*_3, sg*B3galt6*_1: F-CAGGTCCGCAGAAAGGACAT; R-ACCACTCTGTTGTACCTGGC; T_m = 64.9°C

sg*Hs6st*_2, sg*Hs6st*_1: F-ATCCTTTACCAGTACGCGGG; R-AGTGACCCAGGAGGAAGTCT; T_m = 64.7°C

sg*Fgfr2*_1: F-CCTACTTGGGATTGCCAGCA; R-GGTTTCTGCAATCTGGACGC; T_m = 64.7°C

sg*Fgfr2*_3: F-GAAATGGGCCCACTGAGTCA; R-TTCGTGTCTCTCGGTTGTGG; T_m = 64.9°C The resulting amplicons were Sanger sequenced and ICE analysis was performed on the sequences to determine editing efficiency using https://ice.synthego.com/#/

CRISPR-Cas9 screen—Of the genes that were selectively enriched during dormancy in vitro that were encompassed by extracellular matrix-associated gene ontology terms, 95 genes were selected for CRISPR-Cas9 screening in vivo. In addition, positive control sgRNAs targeting genes that are known to be functional during disease progression (pro-

proliferative and pro-survival genes) and negative control sgRNAs that are non-targeting were also included in this library. For the 95 genes, 4–5 sgRNAs were designed for each gene resulting in a total of 509 sgRNAs (Table S1). Additionally, a size-matched library of non-targeting sgRNAs was prepared to aid data analysis.

sgRNAs were designed to target conserved functional domains that display low computationally predicted off-target scores using the GUIDES tool (http://guides.sanjanalab.org/), preferentially picking sgRNAs that have an A/T nucleotide at the 17th position of the sgRNA sequence. This approach maximizes functionally 'null' mutations and circumvents the need for subcloning the cells, thus maintaining their heterogeneity. These sgRNA oligo pools were cloned into the LRG expression vector following which the plasmid pool was amplified, purified, and packaged into lentiviruses. To ensure that *MTB/TAN*-Cas9 cells receive only one sgRNA/cell, the lentiviral library was titered using the GFP selectable marker by flow cytometry (Attune NxT Thermo Fisher) and transduced at an MOI=0.3. Finally, the transduced cells were sorted (MoFlo Astrios, Beckman Coulter Life Sciences) and expanded prior to transplantation in vivo. All steps were performed such that each sgRNA was represented in >1500 cells/sgRNA to maintain coverage and improve the robustness of downstream analyses.

All samples harvested from mouse primary tumors and residual lesions were microdissected under a stereoscope and homogenized for genomic DNA extraction using the Quick-DNA Midiprep Plus kit (Zymo Research, Cat. # D4075). sgRNA inserts were amplified by PCR using the Phusion Flash High Fidelity PCR Master Mix (Thermo Fisher, Cat. # F548), followed by barcode and adapter addition. The libraries were then pooled, mixed with 5% PhiX (Thermo Fisher, Cat. # FC-110-3001) and massively parallel sequenced using the MiSeq reagent kit v3 150 cycle (Illumina, Cat. # MS-102-3001) on the MiSeq Instrument (Illumina).

Histology—When harvesting 3D assays for immunofluorescence, medium was aspirated and immediately fixed in 2% paraformaldehyde (Santa Cruz Biotechnology, Cat. # sc-281692) in PBS for 20 min at room temperature. Cells were then permeabilized in PBS containing 0.5% Triton X-100 for 10 min at room temperature. Wells were then rinsed three times in 1x PBS containing 100mM glycine before proceeding with the immunofluorescence protocol.

For in vitro immunofluorescence/labeling studies, *MTB/TAN*-derived primary tumor cells were cultured on glass coverslips (Bellco Glass Inc., Cat. # 1943–010015A) and treated with 5μM EdU for 2h before harvest. Coverslips containing cells were fixed in 4% paraformaldehyde for 15 min at room temperature. Cells were then permeabilized in PBS containing 0.5% Triton X-100 for 20 min at room temperature followed by wash steps in 3% bovine serum albumin in 1x PBS before proceeding with the immunofluorescence protocol.

For in vivo immunofluorescence/labeling studies, inguinal #4 mammary fat pads containing primary tumors, residual lesions, or recurrent tumors were dissected and fixed overnight in 4% paraformaldehyde overnight at 4°C. Samples were thoroughly washed and dehydrated prior to paraffin embedding and sectioning at 5μm. Slides were prepared by serial deparaffinization and rehydration followed by PBS washes and antigen retrieval in R-Buffer A (Electron Microscopy Sciences, Cat. # 62706–10) or R-Buffer B (Electron Microscopy Sciences, Cat. # 62706–11) using a retriever (Aptum, Cat. # RR2100-EU).

For GAG immunofluorescence, specificity controls were included where tissue sections were treated with Chondroitinase ABC (Amsbio, Cat. # AMS.E1028–02; 10mU/μl at pH 8) or Heparin lyase (Amsbio, Cat. # AMS.HEP-ENZ III-S; 20mU/μl at pH 7) at 37°C for 1h prior to immunofluorescence.

For labeling studies using coverslips for EdU (Thermo Fisher Scientific, Cat. # C10640) or on tissues for EdU (Sigma Aldrich, Cat. # BCK647-IV-IM-S) or TUNEL (Thermo Fisher Scientific, Cat. # C10619) samples were processed as per manufacturer’s instructions prior to immunofluorescence.

Immunofluorescence samples were blocked with 5% goat serum in 1x PBS with mouse-on-mouse block (Vector Laboratories, Cat. # BMK-2202), washed 3 times in 1x PBS and incubated overnight at 4°C with primary antibodies or matched isotype controls diluted in 5% goat serum in 1x PBS with mouse-on-mouse diluent (Vector Laboratories, Cat. # BMK-2202) as follows: Primary antibodies - Rat anti-Ki67 (Thermo Fisher Scientific, Cat. # 14–5698; 1:100), Rabbit anti-Her2 (Cell Signaling, Cat. # 2165; 1:100), Rabbit anti-cleaved Caspase-3 (Cell Signaling, Cat. # 9664; 1:250), Mouse anti-Chondroitin sulfate (Sigma, Cat. # SAB4200696), Mouse anti-Heparan sulfate (Amsbio, Cat. # 370225), Rabbit anti-GFP (Cell Signaling, Cat. # 2956; 1:200), and Mouse anti-GFP (Living Colors, Cat. # 632381; 1:250). Isotype controls - Rat IgG (Thermo Fisher Scientific, Cat. # 02–9602), Rabbit IgG (Thermo Fisher Scientific, Cat. # 02–6102), Mouse IgG2a kappa (Thermo Fisher Scientific, Cat. # 14-4724-82).

After performing 3 washes with 1x PBS, samples were incubated with secondary antibodies at 37°C for 1h as follows: Secondary antibodies - Goat anti-mouse IgG2a Alexa488 (Thermo Fisher Scientific, Cat. # A21131; 1:1000), Goat anti-rabbit IgG Alexa488 (Thermo Fisher Scientific, Cat. # A11034; 1:1000), Goat anti-rabbit IgG Alexa594 (Thermo Fisher Scientific, Cat. # A11012; 1:1000), Goat anti-rat IgG Alexa568 (Thermo Fisher Scientific, Cat. # A11077; 1:1000), Goat anti-mouse IgM Alexa568 (Abcam, Cat. # ab175702; 1:1000).

Following incubation with secondary antibodies, samples were washed 3 times with 1x PBS and incubated with 0.5μg/ml Hoechst 33258 for 10 min at room temperature for nuclear counter staining. Samples were mounted using ProLong gold (Thermo Fisher Scientific, Cat. # P36934) and imaged using a DM 5000B Automated Upright Microscope with a DFC350 FX monochrome digital camera (Leica Microsystems).

Images were quantified using QuPath-0.3.0 open source software. For pixel intensity measurements, a baseline threshold of 8000 pixels was determined based on the background tumor cell fluorescence levels in the heparinase/chondroitin lyase treated controls. Tumor cell boundaries were identified using the GFP signal and heparan sulfate/chondroitin sulfate pixel intensity was measured within a 2μm radius surrounding the tumor cells. Cells with pixel intensity ≤8000 pixels were classified as negative, 8001–16000 pixels were classified as 1+, 16001–24000 pixels were classified as 2+, and 24001–65535 pixels were classified as 3+.

Droplet digital PCR—Microdissected lesions from the in vivo experiments or cells harvested from the in vitro dormancy experiment were processed for genomic DNA isolation using the Quick-DNA Midiprep Plus kit (Zymo Research, Cat. # D4075) kit, or the QIAamp DNA mini kit (Thermo Fisher, Cat. # 51304), respectively. To quantify the numbers of GFP^+ and $mCherry^+$ cells, 50ng of gnomic DNA/sample was added together with the ddPCR Supermix for Probes (Bio-Rad, Cat. # 1863025) and the probes of interest: GFP (Bio-Rad, Cat. # dCNS372378948), mCherry (Bio-Rad, Cat. # dCNS507694046), and the control ApoB (Bio-Rad, Cat. # dMmuCNS4075944696). Droplets were generated using the AutoDG system (Bio-Rad, Cat. # 1864101) followed by PCR under the following conditions: 95°C for 5 min; 40 cycles at 94°C for 30 sec and 60°C for 1 min; 98°C for 10 min. Droplets were analyzed using the QX200 Droplet Reader (Bio-Rad, Cat. # 1864003). Droplet-derived copy numbers were first normalized to ApoB numbers to normalize input and then converted to cell numbers by using the average copy number derived from singly transduced cells, or doubly transduced cells in the combinatorial epistasis experiment.

Western blot—Western blots were performed as described using the following primary antibodies: Rabbit anti-phospho-ERK1/2 (Thr202/Tyr204) (Cell Signaling, Cat. # 9101), Rabbit anti-ERK1/2 (Cell Signaling, Cat. # 9102), Mouse anti-β-Tubulin (BioGenex, Cat. # MU122). Secondary antibodies used were anti-mouse 680LT (LI-COR Biosciences, Cat. # 925–68020) and anti-rabbit 800CW (LI-COR Biosciences, Cat. # 925–32211). Fluorescent signals were detected using the Odyssey detection system (LI-COR Biosciences), and band intensities were quantified using the Image Studio Ver 2.0 software (LI-COR Biosciences).

GRIL-LC/MS—*MTB/TAN*-derived primary tumor cells were cultured for an in vitro dormancy time course as previously described and harvested at baseline (D0), deinduction (D7, D28), and reinduction (D28+) time points. At the time of harvest, cells were kept on ice, thoroughly washed 2 times with ice cold 1x PBS, and gently scraped into 1x PBS. Cells were pelleted down at 1500 rpm and the pellets were flash frozen for further analysis.

Glycan reductive isotope labeling – mass spectrometry (GRIL-LC/MS) – GAG isolation and purification was performed by sonicating the cell pellet in ultrapure distilled water (Invitrogen, Cat. # 10977–015) containing measured amount of protease inhibitor cocktail SetIII, EDTA free (EMD Millipore, Cat. # 539134–1ml) followed by adding equal volume of 2X wash buffer (100mM NaOAc and 400mM NaCl; pH 6). Protein digestion was done using bacterial protease (Sigma, Cat. # P5147) at a concentration of 0.4mg/ml at 37°C overnight (16h). Samples were then loaded on to a Poly-Prep column (Bio-Rad, Cat. # 731–1550) packed with pre-equilibrated DEAE Sephacel gel (Sigma, Cat. # 16505). Columns were washed with 10 bed volume of wash buffer (50mM NaOAc containing 200mM NaCl; pH 6) and bound GAG were eluted using elution buffer (50mM NaOAc containing 1M NaCl; pH 6). Eluted samples were then loaded on to a PD10 desalting column (GE Healthcare, Cat. # 17-0851-01) pre-washed with 10% ethanol. Desalted GAG samples in 10% ethanol were lyophilized and used for further analysis.

Lyophilized samples were resuspended in MilliQ water and distributed for enzymatic digestion of chondroitin sulfate and heparan sulfate at 37°C for 24h using 30mU of Chondroitinase ABC (Amsbio, Cat. # AMS.E1028–02) and 10mU of Heparinase I, II, III (Ibex, Cat. # PN 50–010; PN 50–011; PN 50–012 Heparinase I, II, and III), respectively. Digested samples were spin-filtered using a 3K MWCO spin filtering unit (Pall Life Sciences, Cat. # OD003C34) and tagged with $^{12}C_6$-aniline in the presence of a reducing agent 1M NaCNBr in a mixture of DMSO: HOAc (65:35 v/v). Reductive isotope labeling is performed at 65°C for 45 minutes, followed by incubation at 37°C for 16 h. Samples were mixed with internal standards ($^{13}C_6$-aniline tagged disaccharides from chondroitin sulfate and heparan sulfate respectively) and LC/MS is performed in negative mode using an LTQ-Orbitrap (Thermo Fisher Scientific) mass spectrometer.

Human data analysis—The Glycosaminoglycan biosynthesis 24 gene heparan sulfate/heparin pathway signature and 21 gene chondroitin sulfate/dermatan sulfate signatures were obtained from KEGG (https://www.kegg.jp/kegg-bin/show_pathway?map00534; https://www.kegg.jp/kegg-bin/show_pathway?map00532). The D2.OR 3D dormancy signature was previously described by Prunier *et al*[34]. The 55 gene FGF pathway signature was obtained from the MSigDB database of GSEA (https://www.gsea-msigdb.org/gsea/msigdb/cards/PID_FGF_PATHWAY). For neoadjuvant chemotherapy response analysis, the mean fold change for the time points (post- treatment, surgery) is calculated relative to the pre-treatment (baseline) group in each dataset[21–27] and an ANOVA was performed to obtain p values. For recurrence-free survival analysis, signature scores were calculated for each of the 17 publicly available human breast cancer microarray data sets using the relative expression of the signature genes, as was previously performed[7]. Within each data set, the effect size of the association between gene expression signatures and 10-year relapse-free survival was estimated using the hazard ratio from Cox proportional hazards regression. Effect size estimates were combined across data sets by meta-analysis using the inverse-variance weighting method. The associated p values were obtained using a fixed-effect model.

CN data for METABRIC and TGCA-BRCA[61] was sourced from CBioPortal[62] and took the form of 0 = WT, +1 = gain, +2 = amplification, −1 = loss, −2 = deletion per gene. Clinical data for METABRIC and TCGA-BRCA was sourced from [63] and [62], respectively. Patients that were stage IV at primary diagnosis or had metastasis or recurrence within a month were removed from analysis. Recurrence-free survival was defined as time to and incidence of local recurrence or distant metastasis. Survival coefficient was measured using Cox proportional hazards regression per receptor subtype (HR, HER2, and the combination thereof). When testing across receptor subtypes, HR and HER2 statuses were used as dependent variables.

The raw read counts for the CTC data (GSE113890)[48] were downloaded from NCBI GEO and were normalized using DESeq2. The distribution of expressed genes across all samples ranged from 5000 to greater than 30000 genes. A cutoff of 20000 expressed genes was established to eliminate samples that have low read counts, resulting in the exclusion of 19 patient samples and retention of 8 patient samples in the analysis. A paired t-test was performed to determine the significance of the fold change between the CTC and metastatic groups.

QUANTIFICATION AND STATISTICAL ANALYSIS

For each gene expression signature in each data set, signature scores were calculated as weighted averages of z-score transformed expression data across signature genes, where the z-score transformation was done for each gene across the samples, and the weights are 1 for genes expected to be positively associated with the signature target (e.g. dormancy) and −1 for genes expected to be negatively associated with the signature target.

CRISPR screen sequencing reads were de-convoluted using the unique barcodes associated with each sample and mapped to sgRNA sequences. The abundance of each sgRNA was normalized by the median of ratios (MoR, as implemented in DESeq2), the total read count (T, as % of total), the robust z-score (RZ, with median-centering and MAD-scaling), trimmed mean of M-values (TMM, as implemented in edgeR), and normalized rank (NR, as percentile rank) following which depletion and enrichment scores for each sgRNA were calculated by MAGeCK or the Mann-Whitney U test followed by robust rank aggregation (RRA). False positive rates were controlled for using data from the size matched non-targeting negative control (NC) sgRNA library. In addition to the five normalization methods above, an NC-centric version of MoR was done using negative control gRNAs to calculate scaling factors instead of using all gRNAs, and an NC-centric version of RZ was done using the median and MAD values calculated from negative control gRNAs instead of all gRNAs. Methods with the highest sensitivity were as follows: MAGeCK-RRA-M, MAGeCK-RRA-T, MW-RRA-MoR, MW-RRA-NC_MoR, MW-RRA-NC_RZ, MW-RRA-NR, MW-RRA-RZ, MW-RRA-TMM. Using p values from the 8 tests above, hits called by a larger number of tests were prioritized for further validation.

To determine differences between multiple groups, ANOVA followed by Tukey’s multiple comparisons test was used. In cases where the data were not normally distributed, the Mann-Whitney U test was used to determine statistical significance. Survival curves were subjected to Kaplan-Meier analysis and p values and hazard ratios were calculated using the Mantel-Haenszel method.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Junwei Shi and his laboratory for their advice on developing and performing the CRISPR-Cas9 screen, Dr. Mikala Egeblad for the gift of D2.OR and D2A1 cells, Jianping Wang for histology assistance, Mousumi Paulchakrabarti for GRIL-LC/MS assistance, and Cuyler Luck as well as members of the Chodosh laboratory for their invaluable input that led to the development of techniques used in this study. This project was supported in part by grants from the NIH, the Breast Cancer Research Foundation, and philanthropy from Rhoda Polly Danziger and Michael Danziger to L.A.C. and by a grant from the DOD to A.S.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, and Bray F (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 71, 209–249. 10.3322/caac.21660. [PubMed: 33538338]

2. Pan H, Gray R, Braybrooke J, Davies C, Taylor C, McGale P, Peto R, Pritchard KI, Bergh J, Dowsett M, et al. (2017). 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. N Engl J Med 377, 1836–1846. 10.1056/NEJMoa1701830. [PubMed: 29117498]

3. Pedersen RN, Esen BO, Mellemkjaer L, Christiansen P, Ejlertsen B, Lash TL, Norgaard M, and Cronin-Fenton D (2021). The Incidence of Breast Cancer Recurrence 10–32 Years after Primary Diagnosis. J Natl Cancer Inst. 10.1093/jnci/djab202.

4. Pantel K, Schlimok G, Braun S, Kutter D, Lindemann F, Schaller G, Funke I, Izbicki JR, and Riethmuller G (1993). Differential expression of proliferation-associated molecules in individual micrometastatic carcinoma cells. J Natl Cancer Inst 85, 1419–1424. 10.1093/jnci/85.17.1419. [PubMed: 7688814]

5. Dalla E, Sreekumar A, Aguirre-Ghiso JA, and Chodosh LA (2023). Dormancy in Breast Cancer. Cold Spring Harb Perspect Med. 10.1101/cshperspect.a041331.

6. Moody SE, Sarkisian CJ, Hahn KT, Gunther EJ, Pickup S, Dugan KD, Innocent N, Cardiff RD, Schnall MD, and Chodosh LA (2002). Conditional activation of Neu in the mammary epithelium of transgenic mice results in reversible pulmonary metastasis. Cancer Cell 2, 451–461. 10.1016/s1535-6108(02)00212-x. [PubMed: 12498714]

7. Ruth JR, Pant DK, Pan TC, Seidel HE, Baksh SC, Keister BA, Singh R, Sterner CJ, Bakewell SJ, Moody SE, et al. (2021). Cellular dormancy in minimal residual disease following targeted therapy. Breast Cancer Res 23, 63. 10.1186/s13058-021-01416-9. [PubMed: 34088357]

8. Gunther EJ, Moody SE, Belka GK, Hahn KT, Innocent N, Dugan KD, Cardiff RD, and Chodosh LA (2003). Impact of p53 loss on reversal and recurrence of conditional Wnt-induced tumorigenesis. Genes Dev 17, 488–501. 10.1101/gad.1051603. [PubMed: 12600942]

9. Janghorban M, Yang Y, Zhao N, Hamor C, Nguyen TM, Zhang XH, and Rosen JM (2021). Single Cell Analysis Unveils the Role of the Tumor Immune Microenvironment and Notch Signaling in Dormant Minimal Residual Disease. Cancer Res. 10.1158/0008-5472.CAN-21-1230.

10. Morris VL, Tuck AB, Wilson SM, Percy D, and Chambers AF (1993). Tumor progression and metastasis in murine D2 hyperplastic alveolar nodule mammary tumor cell lines. Clin Exp Metastasis 11, 103–112. 10.1007/BF00880071. [PubMed: 8422701]

11. Morris VL, Koop S, MacDonald IC, Schmidt EE, Grattan M, Percy D, Chambers AF, and Groom AC (1994). Mammary carcinoma cell lines of high and low metastatic potential differ not in extravasation but in subsequent migration and growth. Clin Exp Metastasis 12, 357–367. 10.1007/BF01755879. [PubMed: 7923988]

12. Naumov GN, MacDonald IC, Weinmeister PM, Kerkvliet N, Nadkarni KV, Wilson SM, Morris VL, Groom AC, and Chambers AF (2002). Persistence of solitary mammary carcinoma cells in a secondary site: a possible contributor to dormancy. Cancer Res 62, 2162–2168. [PubMed: 11929839]

13. Ren Q, Khoo WH, Corr AP, Phan TG, Croucher PI, and Stewart SA (2022). Gene expression predicts dormant metastatic breast cancer cell phenotype. Breast Cancer Res 24, 10. 10.1186/s13058-022-01503-5. [PubMed: 35093137]

14. Alvarez JV, Pan TC, Ruth J, Feng Y, Zhou A, Pant D, Grimley JS, Wandless TJ, Demichele A, Investigators IST, and Chodosh LA (2013). Par-4 downregulation promotes breast cancer recurrence by preventing multinucleation following targeted therapy. Cancer Cell 24, 30–44. 10.1016/j.ccr.2013.05.007. [PubMed: 23770012]

15. Feng Y, Pan TC, Pant DK, Chakrabarti KR, Alvarez JV, Ruth JR, and Chodosh LA (2014). SPSB1 promotes breast cancer recurrence by potentiating c-MET signaling. Cancer Discov 4, 790–803. 10.1158/2159-8290.CD-13-0548. [PubMed: 24786206]

16. Payne AW, Pant DK, Pan TC, and Chodosh LA (2014). Ceramide kinase promotes tumor cell survival and mammary tumor recurrence. Cancer Res 74, 6352–6363. 10.1158/0008-5472.CAN-14-1292. [PubMed: 25164007]

17. D'Cruz CM, Gunther EJ, Boxer RB, Hartman JL, Sintasath L, Moody SE, Cox JD, Ha SI, Belka GK, Golant A, et al. (2001). c-MYC induces mammary tumorigenesis by means of a preferred pathway involving spontaneous Kras2 mutations. Nat Med 7, 235–239. 10.1038/84691. [PubMed: 11175856]

18. Gunther EJ, Belka GK, Wertheim GB, Wang J, Hartman JL, Boxer RB, and Chodosh LA (2002). A novel doxycycline-inducible system for the transgenic analysis of mammary gland biology. FASEB J 16, 283–292. 10.1096/fj.01-0551com. [PubMed: 11874978]
19. Alvarez JV, Belka GK, Pan TC, Chen CC, Blankemeyer E, Alavi A, Karp JS, and Chodosh LA (2014). Oncogene pathway activation in mammary tumors dictates FDG-PET uptake. Cancer Res 74, 7583–7598. 10.1158/0008-5472.CAN-14-1235. [PubMed: 25239452]
20. Abravanel DL, Belka GK, Pan TC, Pant DK, Collins MA, Sterner CJ, and Chodosh LA (2015). Notch promotes recurrence of dormant tumor cells following HER2/neu-targeted therapy. J Clin Invest 125, 2484–2496. 10.1172/JCI74883. [PubMed: 25961456]
21. Stickeler E, Pils D, Klar M, Orlowsk-Volk M, Zur Hausen A, Jager M, Watermann D, Gitsch G, Zeillinger R, and Tempfer CB (2011). Basal-like molecular subtype and HER4 up-regulation and response to neoadjuvant chemotherapy in breast cancer. Oncol Rep 26, 1037–1045. 10.3892/or.2011.1392. [PubMed: 21769435]
22. Turnbull AK, Arthur LM, Renshaw L, Larionov AA, Kay C, Dunbier AK, Thomas JS, Dowsett M, Sims AH, and Dixon JM (2015). Accurate Prediction and Validation of Response to Endocrine Therapy in Breast Cancer. J Clin Oncol 33, 2270–2278. 10.1200/JCO.2014.57.8963. [PubMed: 26033813]
23. Arthur LM, Turnbull AK, Webber VL, Larionov AA, Renshaw L, Kay C, Thomas JS, Dixon JM, and Sims AH (2014). Molecular changes in lobular breast cancers in response to endocrine therapy. Cancer Res 74, 5371–5376. 10.1158/0008-5472.CAN-14-0620. [PubMed: 25100562]
24. Miller WR, Larionov AA, Renshaw L, Anderson TJ, White S, Murray J, Murray E, Hampton G, Walker JR, Ho S, et al. (2007). Changes in breast cancer transcriptional profiles after treatment with the aromatase inhibitor, letrozole. Pharmacogenet Genomics 17, 813–826. 10.1097/FPC.0b013e32820b853a. [PubMed: 17885619]
25. Korde LA, Lusa L, McShane L, Lebowitz PF, Lukes L, Camphausen K, Parker JS, Swain SM, Hunter K, and Zujewski JA (2010). Gene expression pathway analysis to predict response to neoadjuvant docetaxel and capecitabine for breast cancer. Breast Cancer Res Treat 119, 685–699. 10.1007/s10549-009-0651-3. [PubMed: 20012355]
26. Gruosso T, Mieulet V, Cardon M, Bourachot B, Kieffer Y, Devun F, Dubois T, Dutreix M, Vincent-Salomon A, Miller KM, and Mechta-Grigoriou F (2016). Chronic oxidative stress promotes H2AX protein degradation and enhances chemosensitivity in breast cancer patients. EMBO Mol Med 8, 527–549. 10.15252/emmm.201505891. [PubMed: 27006338]
27. Gonzalez-Angulo AM, Iwamoto T, Liu S, Chen H, Do KA, Hortobagyi GN, Mills GB, Meric-Bernstam F, Symmans WF, and Pusztai L (2012). Gene expression, molecular class changes, and pathway analysis after neoadjuvant systemic therapy for breast cancer. Clin Cancer Res 18, 1109–1119. 10.1158/1078-0432.CCR-11-2762. [PubMed: 22235097]
28. Di Martino JS, Nobre AR, Mondal C, Taha I, Farias EF, Fertig EJ, Naba A, Aguirre-Ghiso JA, and Bravo-Cordero JJ (2022). A tumor-derived type III collagen-rich ECM niche regulates tumor cell dormancy. Nature cancer 3, 90–107. 10.1038/s43018-021-00291-9. [PubMed: 35121989]
29. Bai X, Zhou D, Brown JR, Crawford BE, Hennet T, and Esko JD (2001). Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). J Biol Chem 276, 48189–48195. 10.1074/jbc.M107339200. [PubMed: 11551958]
30. Delbaere S, De Clercq A, Mizumoto S, Noborn F, Bek JW, Alluyn L, Gistelinck C, Syx D, Salmon PL, Coucke PJ, et al. (2020). b3galt6 Knock-Out Zebrafish Recapitulate beta3GalT6-Deficiency Disorders in Human and Reveal a Trisaccharide Proteoglycan Linkage Region. Front Cell Dev Biol 8, 597857. 10.3389/fcell.2020.597857. [PubMed: 33363150]
31. Conant D, Hsiau T, Rossi N, Oki J, Maures T, Waite K, Yang J, Joshi S, Kelso R, Holden K, et al. (2022). Inference of CRISPR Edits from Sanger Trace Data. CRISPR J 5, 123–130. 10.1089/crispr.2021.0113. [PubMed: 35119294]
32. Barkan D, El Touny LH, Michalowski AM, Smith JA, Chu I, Davis AS, Webster JD, Hoover S, Simpson RM, Gauldie J, and Green JE (2010). Metastatic growth from dormant cells induced by a col-I-enriched fibrotic environment. Cancer Res 70, 5706–5716. 10.1158/0008-5472.CAN-09-2356. [PubMed: 20570886]

33. Barkan D, Kleinman H, Simmons JL, Asmussen H, Kamaraju AK, Hoenorhoff MJ, Liu ZY, Costes SV, Cho EH, Lockett S, et al. (2008). Inhibition of metastatic outgrowth from single dormant tumor cells by targeting the cytoskeleton. Cancer Res 68, 6241–6250. 10.1158/0008-5472.CAN-07-6849. [PubMed: 18676848]

34. Prunier C, Alay A, van Dijk M, Ammerlaan KL, van Gelderen S, Marvin DL, Teunisse A, Slieker RC, Szuhai K, Jochemsen AG, et al. (2021). Breast cancer dormancy is associated with a 4NG1 state and not senescence. NPJ Breast Cancer 7, 140. 10.1038/s41523-021-00347-0. [PubMed: 34707097]

35. Prydz K, and Dalen KT (2000). Synthesis and sorting of proteoglycans. J Cell Sci 113 Pt 2, 193–205. [PubMed: 10633071]

36. Rapraeger AC, Krufka A, and Olwin BB (1991). Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. Science 252, 1705–1708. 10.1126/science.1646484. [PubMed: 1646484]

37. Johnson CE, Crawford BE, Stavridis M, Ten Dam G, Wat AL, Rushton G, Ward CM, Wilson V, van Kuppevelt TH, Esko JD, et al. (2007). Essential alterations of heparan sulfate during the differentiation of embryonic stem cells to Sox1-enhanced green fluorescent protein-expressing neural progenitor cells. Stem Cells 25, 1913–1923. 10.1634/stemcells.2006-0445. [PubMed: 17464092]

38. Ashikari-Hada S, Habuchi H, Sugaya N, Kobayashi T, and Kimata K (2009). Specific inhibition of FGF-2 signaling with 2-O-sulfated octasaccharides of heparan sulfate. Glycobiology 19, 644–654. 10.1093/glycob/cwp031. [PubMed: 19254961]

39. Merry CLR, Lindahl U, Couchman J, and Esko JD (2022). Proteoglycans and Sulfated Glycosaminoglycans. In Essentials of Glycobiology, th, Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Mohnen D, Kinoshita T, Packer NH, et al., eds. pp. 217–232. 10.1101/glycobiology.4e.17.

40. Marques C, Reis CA, Vives RR, and Magalhaes A (2021). Heparan Sulfate Biosynthesis and Sulfation Profiles as Modulators of Cancer Signalling and Progression. Front Oncol 11, 778752. 10.3389/fonc.2021.778752. [PubMed: 34858858]

41. Lawrence R, Olson SK, Steele RE, Wang L, Warrior R, Cummings RD, and Esko JD (2008). Evolutionary differences in glycosaminoglycan fine structure detected by quantitative glycan reductive isotope labeling. J Biol Chem 283, 33674–33684. 10.1074/jbc.M804288200. [PubMed: 18818196]

42. Lawrence R, Lu H, Rosenberg RD, Esko JD, and Zhang L (2008). Disaccharide structure code for the easy representation of constituent oligosaccharides from glycosaminoglycans. Nat Methods 5, 291–292. 10.1038/nmeth0408-291. [PubMed: 18376390]

43. Sugaya N, Habuchi H, Nagai N, Ashikari-Hada S, and Kimata K (2008). 6-O-sulfation of heparan sulfate differentially regulates various fibroblast growth factor-dependent signalings in culture. J Biol Chem 283, 10366–10376. 10.1074/jbc.M705948200. [PubMed: 18281280]

44. Qiu H, Shi S, Yue J, Xin M, Nairn AV, Lin L, Liu X, Li G, Archer-Hartmann SA, Dela Rosa M, et al. (2018). A mutant-cell library for systematic analysis of heparan sulfate structure-function relationships. Nat Methods 15, 889–899. 10.1038/s41592-018-0189-6. [PubMed: 30377379]

45. Weiss RJ, Spahn PN, Chiang AWT, Liu Q, Li J, Hamill KM, Rother S, Clausen TM, Hoeksema MA, Timm BM, et al. (2021). Genome-wide screens uncover KDM2B as a modifier of protein binding to heparan sulfate. Nat Chem Biol 17, 684–692. 10.1038/s41589-021-00776-9. [PubMed: 33846619]

46. Ornitz DM, and Itoh N (2015). The Fibroblast Growth Factor signaling pathway. Wiley Interdiscip Rev Dev Biol 4, 215–266. 10.1002/wdev.176. [PubMed: 25772309]

47. Huang ML, Smith RA, Trieger GW, and Godula K (2014). Glycocalyx remodeling with proteoglycan mimetics promotes neural specification in embryonic stem cells. J Am Chem Soc 136, 10565–10568. 10.1021/ja505012a. [PubMed: 25019314]

48. Ring A, Campo D, Porras TB, Kaur P, Forte VA, Tripathy D, Lu J, Kang I, Press MF, Jeong YJ, et al. (2022). Circulating Tumor Cell Transcriptomics as Biopsy Surrogates in Metastatic Breast Cancer. Ann Surg Oncol 29, 2882–2894. 10.1245/s10434-021-11135-2. [PubMed: 35000083]

49. Bayne LJ, Nivar I, Goodspeed B, Deluca SE, Wileyto P, Shih NNC, Nayak A, Feldman MD, Edwards J, Fox K, et al. (2021). Identifying breast cancer survivors with dormant disseminated tumor cells: The PENN-SURMOUNT screening study [abstract]. Cancer Res 81 (4).

50. Bayne LJ, Nivar I, Goodspeed B, Wileyto P, Savage J, Shih NNC, Feldman MD, Edwards J, Clark AS, Fox KR, et al. (2018). Detection and targeting of minimal residual disease in breast cancer to reduce recurrence: The PENN-SURMOUNT and CLEVER trials [abstract]. Cancer Res 78 (4).

51. Malfait F, Kariminejad A, Van Damme T, Gauche C, Syx D, Merhi-Soussi F, Gulberti S, Symoens S, Vanhauwaert S, Willaert A, et al. (2013). Defective initiation of glycosaminoglycan synthesis due to B3GALT6 mutations causes a pleiotropic Ehlers-Danlos-syndrome-like connective tissue disorder. Am J Hum Genet 92, 935–945. 10.1016/j.ajhg.2013.04.016. [PubMed: 23664118]

52. Nakajima M, Mizumoto S, Miyake N, Kogawa R, Iida A, Ito H, Kitoh H, Hirayama A, Mitsubuchi H, Miyazaki O, et al. (2013). Mutations in B3GALT6, which encodes a glycosaminoglycan linker region enzyme, cause a spectrum of skeletal and connective tissue disorders. Am J Hum Genet 92, 927–934. 10.1016/j.ajhg.2013.04.003. [PubMed: 23664117]

53. Turner N, and Grose R (2010). Fibroblast growth factor signalling: from development to cancer. Nat Rev Cancer 10, 116–129. 10.1038/nrc2780. [PubMed: 20094046]

54. Bado IL, Zhang W, Hu J, Xu Z, Wang H, Sarkar P, Li L, Wan YW, Liu J, Wu W, et al. (2021). The bone microenvironment increases phenotypic plasticity of ER(+) breast cancer cells. Dev Cell 56, 1100–1117 e1109. 10.1016/j.devcel.2021.03.008. [PubMed: 33878299]

55. Barrios J, and Wieder R (2009). Dual FGF-2 and intergrin alpha5beta1 signaling mediate GRAF-induced RhoA inactivation in a model of breast cancer dormancy. Cancer Microenviron 2, 33–47. 10.1007/s12307-009-0019-6. [PubMed: 19308677]

56. Nobre AR, Dalla E, Yang J, Huang X, Wullkopf L, Risson E, Razghandi P, Anton ML, Zheng W, Seoane JA, et al. (2022). ZFP281 drives a mesenchymal-like dormancy program in early disseminated breast cancer cells that prevents metastatic outgrowth in the lung. Nat Cancer. 10.1038/s43018-022-00424-8.

57. Subbiah V, Sahai V, Maglic D, Bruderek K, Toure BB, Zhao S, Valverde R, O'Hearn PJ, Moustakas DT, Schonherr H, et al. (2023). RLY-4008, the first highly selective FGFR2 inhibitor with activity across FGFR2 alterations and resistance mutations. Cancer Discov. 10.1158/2159-8290.CD-23-0475.

58. Schram AM, Kamath SD, El-Khoueiry AB, Borad MJ, Mody K, Mahipal A, Goyal L, Sahai V, Schmidt-Kittler O, Shen J, et al. (2021). First-in-human study of highly selective FGFR2 inhibitor, RLY-4008, in patients with intrahepatic cholangiocarcinoma and other advanced solid tumors. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 10.1200/JCO.2021.39.15_suppl.TPS4165 Journal of Clinical Oncology 39, no. 15_suppl.

59. Moody SE, Perez D, Pan TC, Sarkisian CJ, Portocarrero CP, Sterner CJ, Notorfrancesco KL, Cardiff RD, and Chodosh LA (2005). The transcriptional repressor Snail promotes mammary tumor recurrence. Cancer Cell 8, 197–209. 10.1016/j.ccr.2005.07.009. [PubMed: 16169465]

60. Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, and Vakoc CR (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nat Biotechnol 33, 661–667. 10.1038/nbt.3235. [PubMed: 25961408]

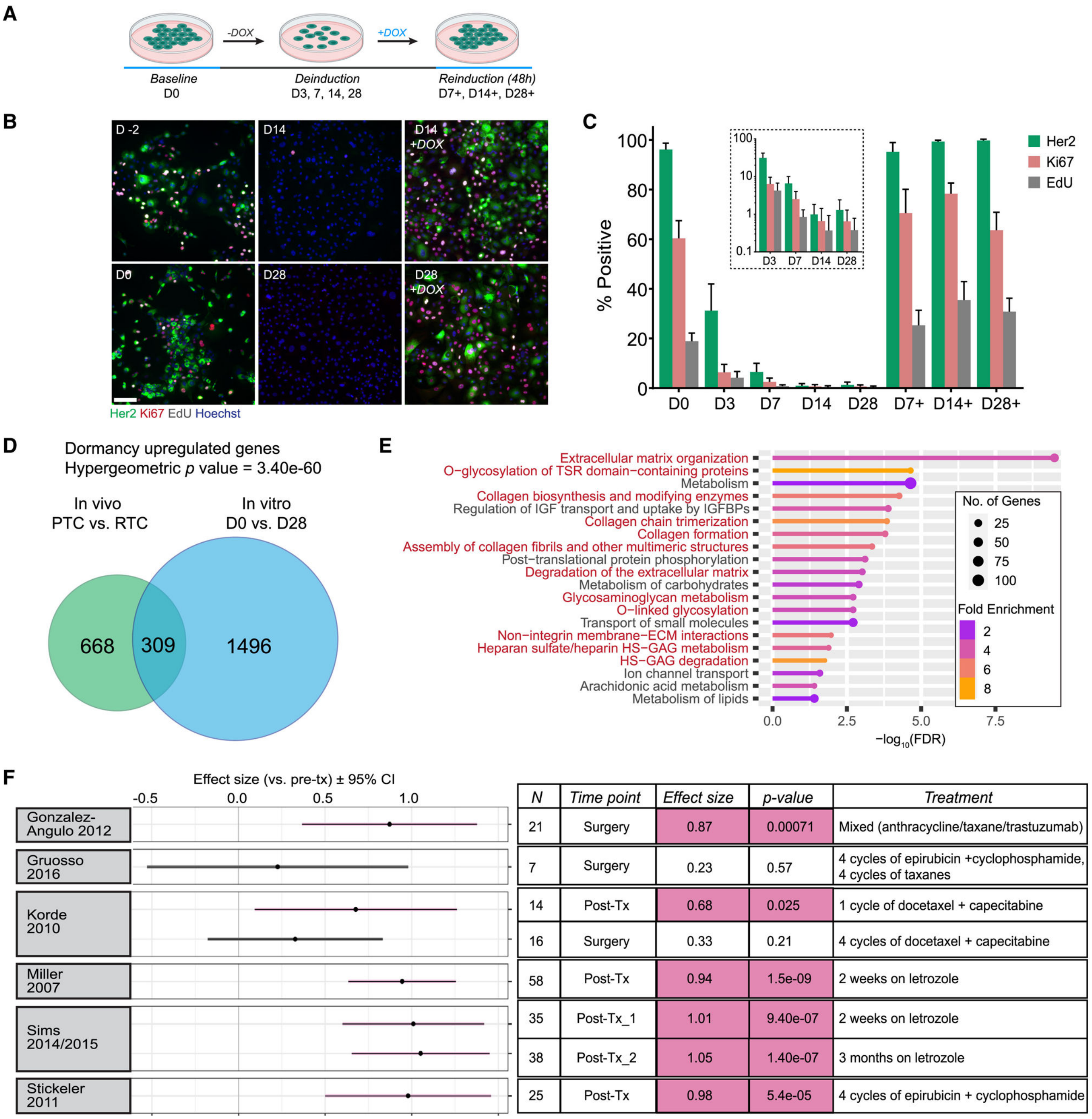
61. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C, et al. (2015). Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. Cell 163, 506–519. 10.1016/j.cell.2015.09.033. [PubMed: 26451490]

62. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery 2, 401–404. 10.1158/2159-8290.Cd-12-0095. [PubMed: 22588877]

63. Rueda OM, Sammut S-J, Seoane JA, Chin S-F, Caswell-Jin JL, Callari M, Batra R, Pereira B, Bruna A, Ali HR, et al. (2019). Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. Nature 567, 399–404. 10.1038/s41586-019-1007-8. [PubMed: 30867590]

Highlights

- Dormant residual tumor cells (RTCs) upregulate heparan sulfate (HS) biosynthesis
- CRISPR-Cas9 screen identifies B3galt6 as a regulator of RTC survival and recurrence
- B3galt6 is required for glycosaminoglycan-protein linkage to generate proteoglycans
- Dormant RTCs upregulate B3galt6/Hs6st1/Fgf1/Fgfr2 signaling to maintain survival



N	Time point	Effect size	p-value	Treatment
21	Surgery	0.87	0.00071	Mixed (anthracycline/taxane/trastuzumab)
7	Surgery	0.23	0.57	4 cycles of epirubicin +cyclophosphamide, 4 cycles of taxanes
14	Post-Tx	0.68	0.025	1 cycle of docetaxel + capecitabine
16	Surgery	0.33	0.21	4 cycles of docetaxel + capecitabine
58	Post-Tx	0.94	1.5e-09	2 weeks on letrozole
35	Post-Tx_1	1.01	9.40e-07	2 weeks on letrozole
38	Post-Tx_2	1.05	1.40e-07	3 months on letrozole
25	Post-Tx	0.98	5.4e-05	4 cycles of epirubicin + cyclophosphamide

Fig. 1: Dormant tumor cells display cell-autonomous upregulation of ECM-related genes following therapy

A. Experimental schematic indicating time points for the in vitro dormancy (IVD) assay using *MTB/TAN* cells. **B.** Immunofluorescence for Her2 (green), Ki67 (red), and EdU (grey) with Hoechst nuclear staining (blue) in IVD. Scale bar=100μm. **C.** Quantification of **(B)** represented as mean ± standard deviation (SD). Inset displays the percentages at deinduction time points on a $\log_{10}$ y-axis. **D.** Hypergeometric test and Venn diagram depicting the overlap between the in vivo and IVD-derived upregulated genes. **E.** Top 20 Reactome gene ontology terms for the overlapping upregulated gene set. ECM-associated categories highlighted in red. **F.** Core RTC signature enrichment following neoadjuvant therapies vs. pre-treatment (pre-tx) samples across 6 patient datasets measured by effect size (mean of pair-wise difference in signature scores). Significant effects are highlighted in pink.

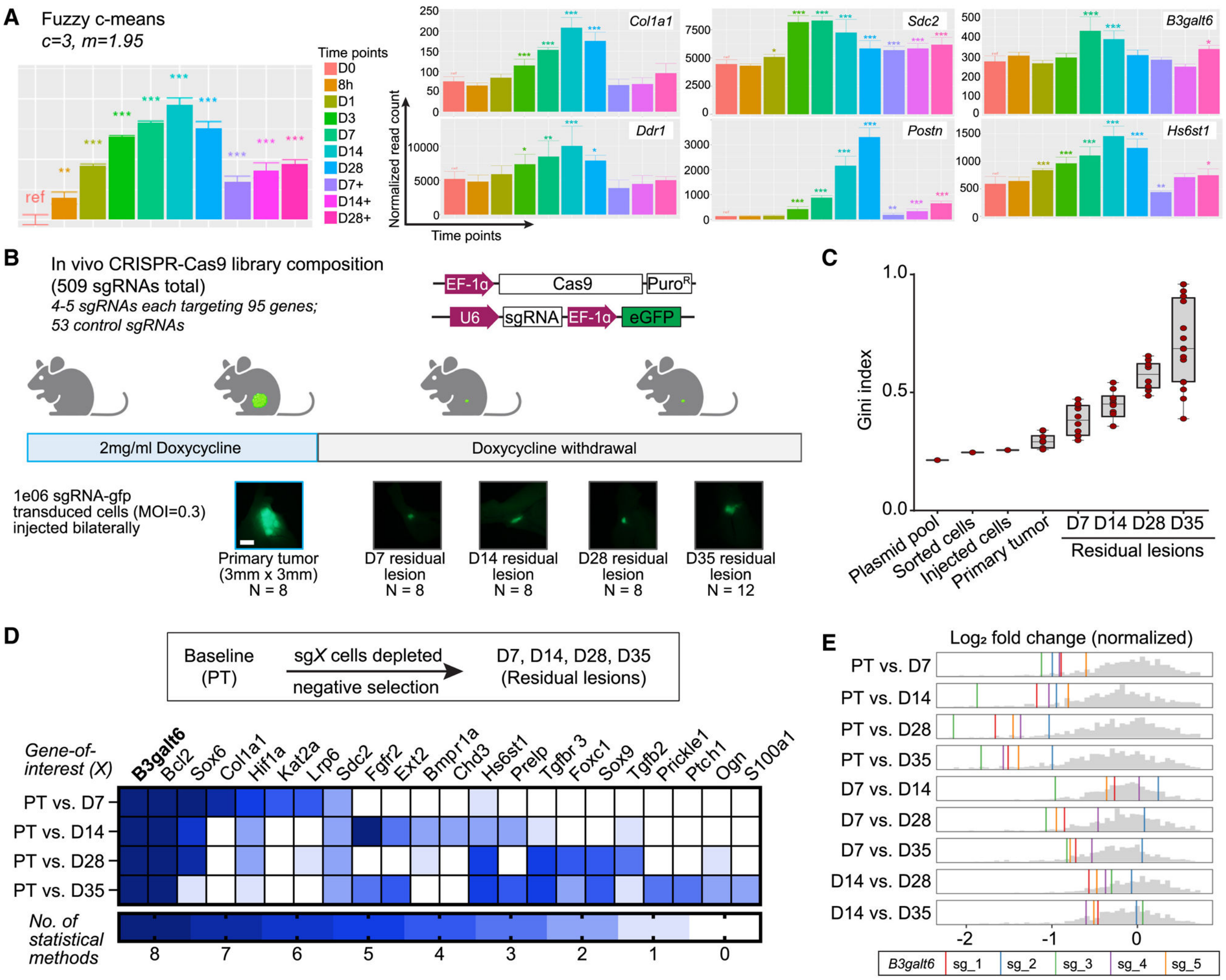


Fig. 2: An ECM-focused loss-of-function screen identifies *B3galt6* as a regulator of RTC fitness in vivo

A. Fuzzy c-means clustering using cluster # (c)=3, fuzzifier (m)=1.95 identifies one cluster that displays reversible, dormancy-dependent upregulation of genes at 8h, D1, D3, D7, D14, D28 (*deinduction*) time points. 6 genes (of 3450) in this cluster are indicated. Asterisks indicate significant changes in normalized read counts vs. D0 (*baseline*) $*p<0.05$, $**p<0.01$, $***p<0.001$, $****p<0.0001$. **B.** In vivo CRISPR-Cas9 screen schematic. Stereoscope images of representative lesions are shown. Scale bar=2mm. **C.** Gini index for heterogeneity at sequential time points assayed in the screen. Data are represented as median ± range. **D.** Target identification criterion and list of CRISPR-Cas9 screen depletion hits identified by 2–8 statistical methods used for calling hits. Top hit *B3galt6* is highlighted (bold). **E.** Log_2 fold change (normalized) gene-level effect size plots for *B3galt6* sgRNAs in different pairwise comparisons using PT, D7, or D14 as baselines. Grey histogram depicts the background distribution of sgRNAs in the CRISPR-Cas9 screen, colored vertical lines depict each sgRNA targeting *B3galt6*.

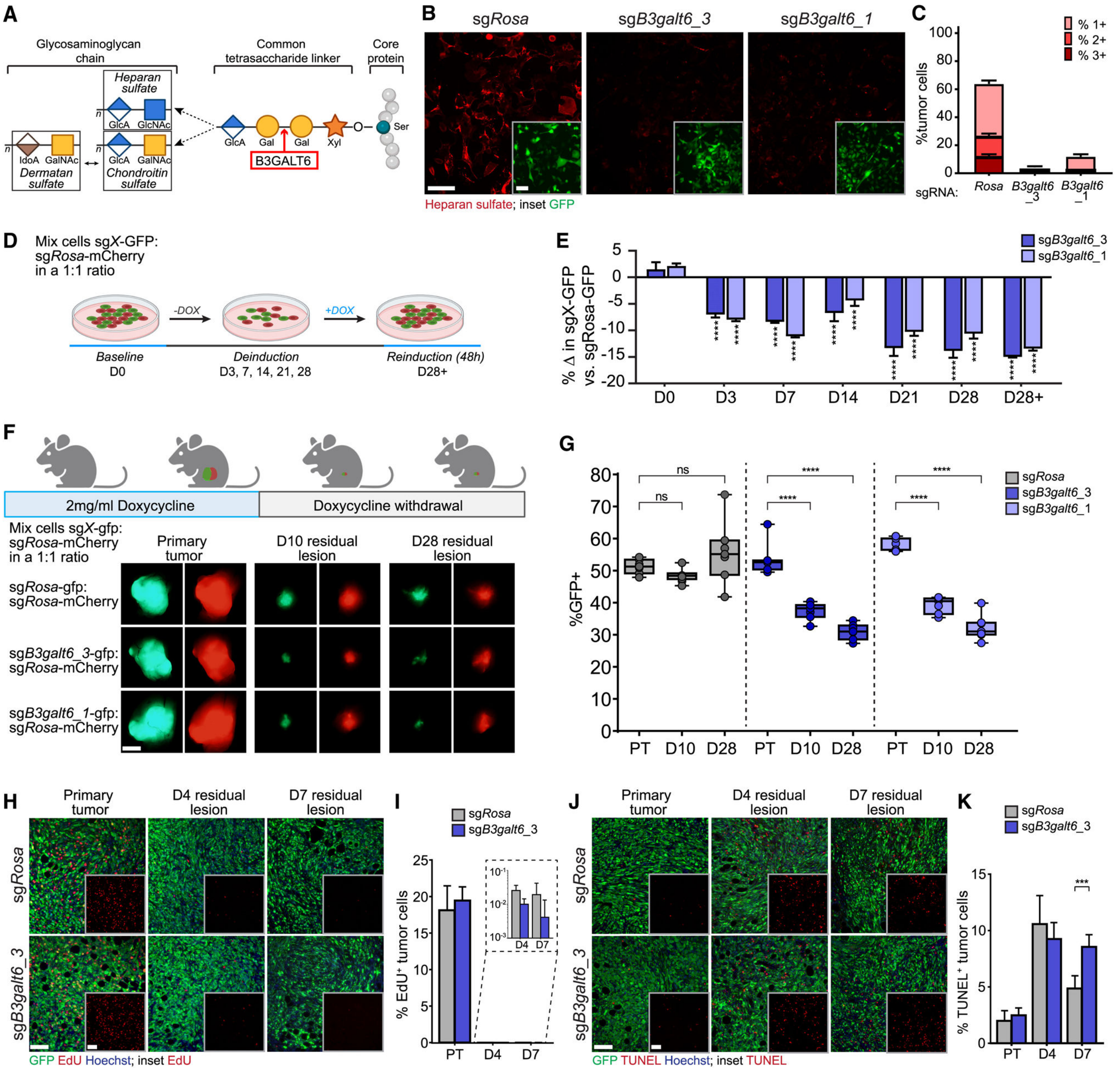


Fig. 3: B3GALT6 promotes RTC survival during dormancy

A. B3GALT6 (red box) function in tetrasaccharide linker synthesis and proteoglycan assembly. Typical disaccharide repeat units in heparan sulfate, chondroitin sulfate, and dermatan sulfate are indicated. Ser=Serine, Xyl=Xylose, Gal=Galactose, GlcA=Glucuronic acid, GlcNAc=*N*-Acetylglucosamine, GalNAc=*N*-Acetylgalactosamine, IdoA=Iduronic acid. **B.** Immunofluorescence and **C.** distribution of pixel intensities for heparan sulfate (red) in Her2-dependent-Cas9 cells transduced with sg*Rosa*, sg*B3galt6*_3, or sg*B3galt6*_1 sgRNAs (green). Scale bar=100μm. 1+, 2+, 3+ refer to pixel intensity bins corresponding to 8001–16000, 16001–24000, and 24001–65535 pixels, respectively. Data are represented as mean ± standard error of mean (SEM). **D.** IVD competition assay schematic. **E.** ddPCR data quantifying percentage change in sg*B3galt6*_3 (dark blue) and sg*B3galt6*_1 (light blue)

GFP+ cells normalized to sg*Rosa* (grey) cell numbers. Data are represented as mean ± SD. n=3 biological replicates/group. ****p<0.0001. **F.** Schematic for in vivo competition assay with stereoscope images of representative lesions harvested. Scale bar=2mm. **G.** ddPCR data quantifying GFP+ cell percentage in sgRosa (grey), sg*B3galt6*_3 (dark blue), and sg*B3galt6*_1 (light blue) groups. Data are represented as median ± range. ns=non-significant, ****p<0.0001. n=3 PT, 7 D10 RL, 7 D28 RL/group. **H-K.** Immunofluorescence **(H)** and quantification **(I)** for EdU (red) or for TUNEL (red) **(J, K)** in Her2-dependent-Cas9 sg*Rosa* and sg*B3galt6*_3 (green) PTs, D4, and D7 residual lesions (RLs). Scale bar=100μm. Quantification in the sg*Rosa* (grey) and sg*B3galt6*_3 (dark blue) groups is represented as mean ± SD. n=6 PT, 8 D4 RL, 10 D7 RL/group. ***p<0.001.

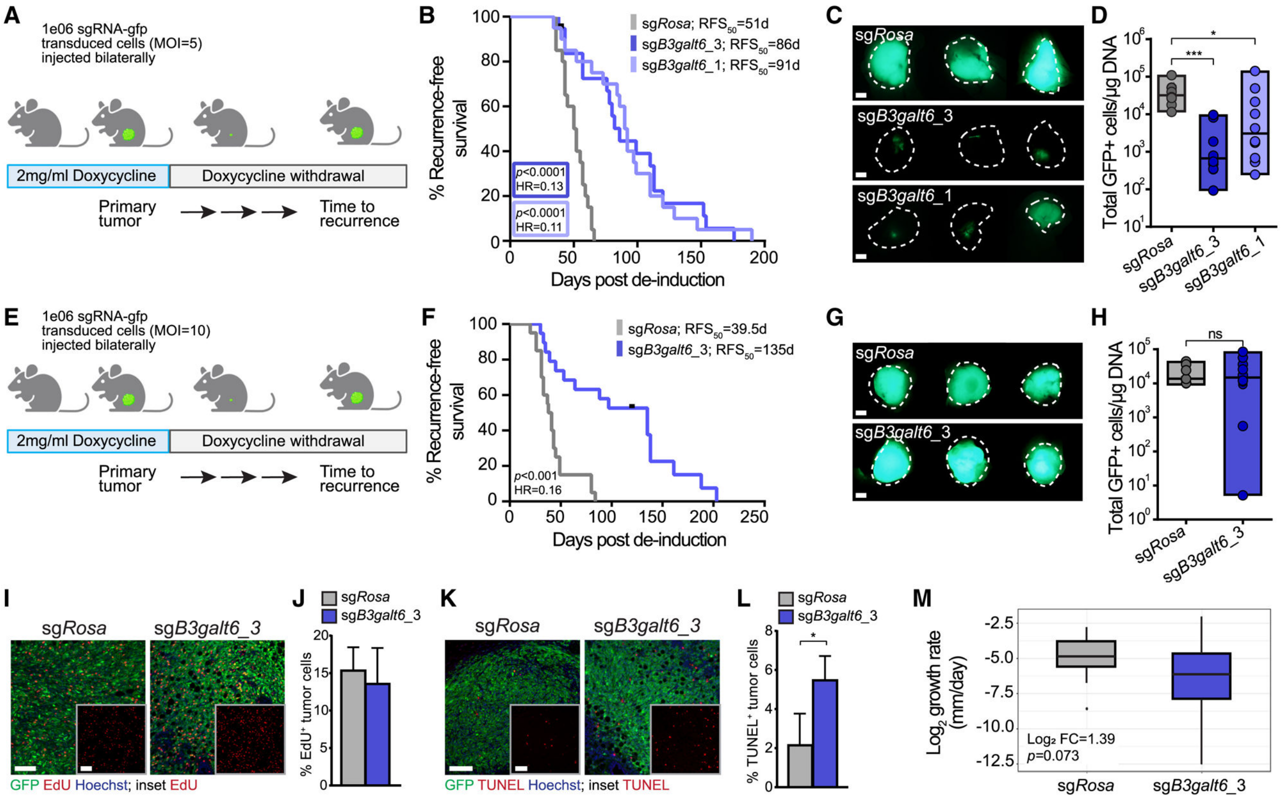


Fig. 4: B3GALT6 promotes recurrence following dormancy

A. Recurrence-free survival assay using multiplicity of infection (MOI)=5. **B.** Kaplan-Meier analysis of recurrence-free survival for sg*Rosa* (grey), sg*B3galt6*_3 (dark blue), and sg*B3galt6*_1 (light blue) groups. n=20 mice/group. RFS_{50}=median time-to-recurrence. **C.** Stereoscope images of representative recurrences. Dotted white lines represent tumor edges identified in corresponding bright field images. Scale bar=2mm. **D.** Quantification of GFP+ cells in recurrent tumors as measured by ddPCR for sg*Rosa* (grey), sg*B3galt6*_3 (dark blue), and sg*B3galt6*_1 (light blue) groups. Data are represented as median ± range. *$p<0.05$, ***$p<0.001$. **E.** Recurrence-free survival assay using MOI=10. **F.** Kaplan-Meier analysis of recurrence-free survival for sg*Rosa* (grey) and sg*B3galt6*_3 (dark blue) groups. n=20 mice/group. RFS_{50}=median time-to-recurrence. **G.** Stereoscope images of representative recurrences. Scale bar=2mm. **H.** Quantification of GFP+ cells measured by ddPCR for sg*Rosa* (grey) and sg*B3galt6*_3 (dark blue) recurrences. Data are represented as median ± range. ns=non-significant. **I-L.** Immunofluorescence **(I)** and quantification **(J)** for EdU (red) or for TUNEL (red) **(K, L)** in Her2-dependent-Cas9 sg*Rosa* and sg*B3galt6*_3 (green) recurrences. sg*Rosa* recurrences were harvested 41–83d post-de-induction, sg*B3galt6*_3 recurrences were harvested 53–139d post-de-induction. Scale bar=100μm. Quantification in sg*Rosa* (grey) and sg*B3galt6*_3 (dark blue) groups is represented as mean ± SD. *$p<0.05$. **M.** Log_2 growth rate of sg*Rosa* (grey) and sg*B3galt6*_3 (dark blue) recurrent tumors. Data are represented as median ± interquartile range.

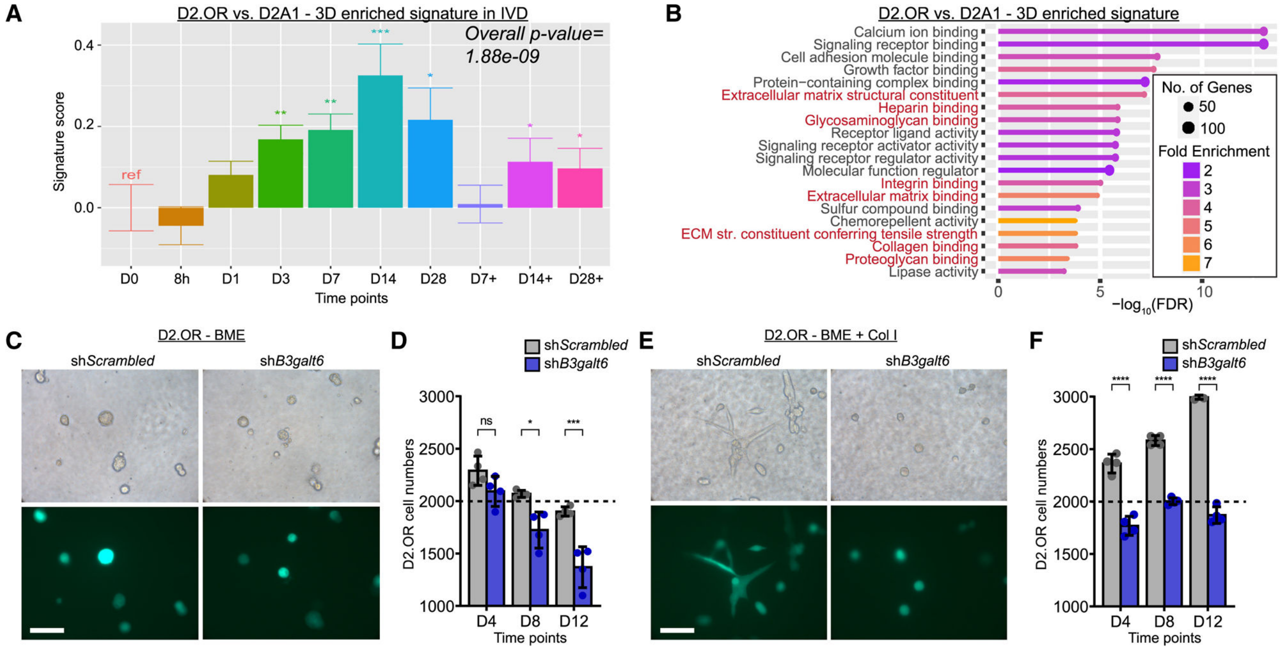


Fig. 5: B3GALT6 promotes tumor cell survival and outgrowth in microenvironment-induced models of dormancy

A. Application of a gene expression signature derived from D2.OR (indolent) cells vs. D2A1 (aggressive) cells in 3D to IVD temporal profiling of Her2-dependent tumor cells. Asterisks indicate significant changes in normalized read counts vs. D0 (*baseline*) *$p<0.05$, **$p<0.01$, ***$p<0.001$. **B.** Top 20 gene ontology terms for the upregulated set of genes selectively enriched in D2.OR cells in 3D. ECM-associated categories are highlighted in red. **C.** Brightfield and fluorescence images of sh*Scrambled* and sh*B3galt6* D2.OR cells grown in 3D on basement membrane extract (BME) and **D.** viable cell numbers measured at D4, D8, and D12 time points. Dotted line indicates cell number at D0. Scale bar=100μm. Data are represented as mean ± SD. n=4 replicates/group. ns=non-significant, *$p<0.05$, ****$p<0.0001$. **E.** Brightfield and fluorescence images of sh*Scrambled* and sh*B3galt6* D2.OR cells grown in 3D on BME + collagen I (Col I) and **F.** viable cell numbers measured at D4, D8, and D12 time points. Dotted line indicates cell number at D0. Scale bar=100μm. Data are represented as mean ± SD. n=4 replicates/group. ns=non-significant, *$p<0.05$, ****$p<0.0001$.

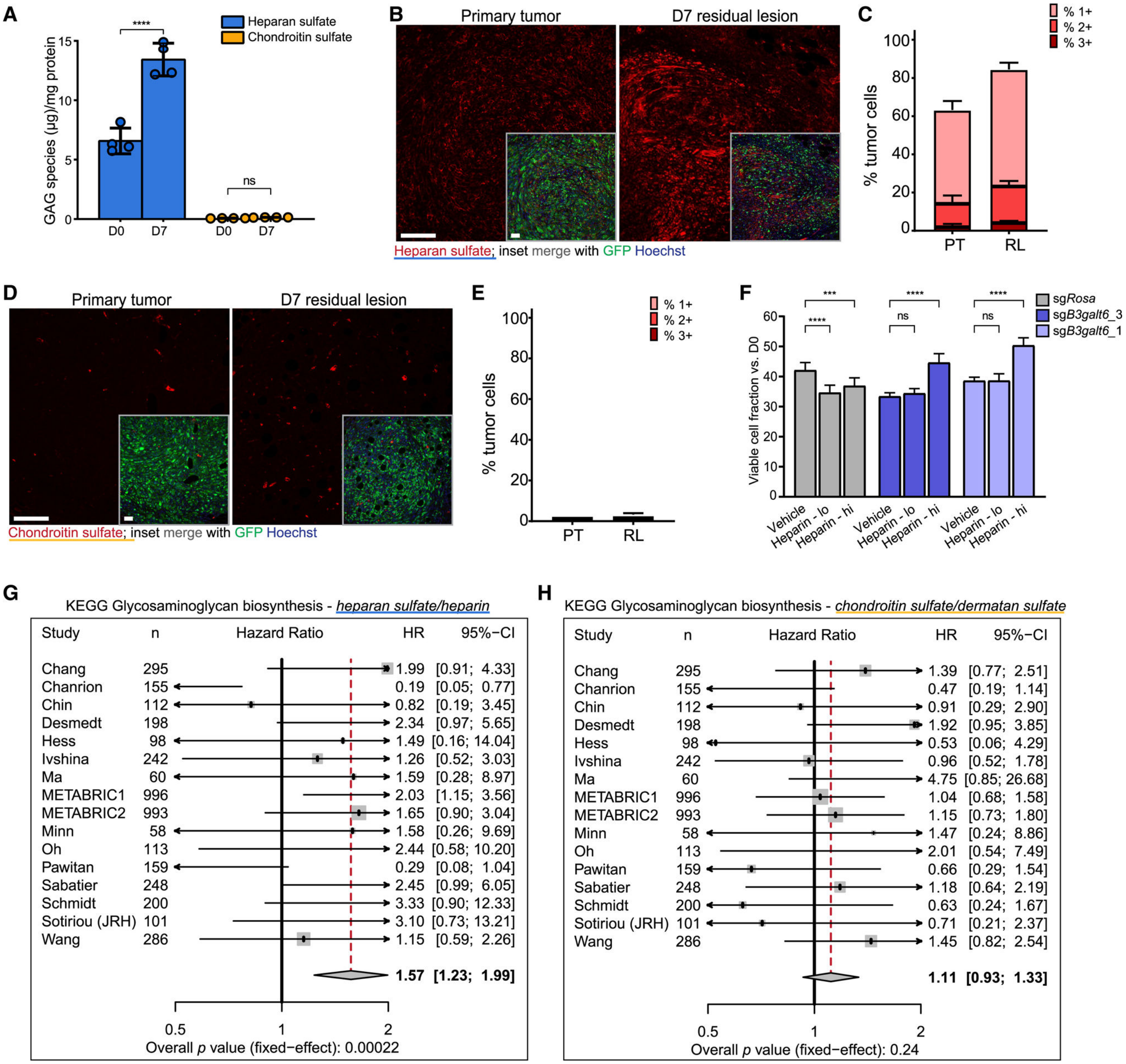


Fig. 6: Heparan sulfate synthesis is upregulated during dormancy and associated with worse recurrence-free survival in patients with breast cancer

A. Liquid chromatography/mass spectrometry (LC/MS) analysis of heparan sulfate (blue) and chondroitin sulfate (yellow) GAGs performed on D0 and D7 in vitro samples. Data are represented as mean ± SD. n=4 biological replicates/group. ns=non-significant, ****p<0.0001. **B-E.** Immunofluorescence and distribution of pixel intensities for heparan sulfate **(B, C)** and chondroitin sulfate **(D, E)** (red) in PTs and D7 RLs derived from Her2-dependent-Cas9 cells with sg*Rosa* (green). Scale bar=100µm. 1+, 2+, 3+ refer to pixel intensity bins corresponding to 8001–16000, 16001–24000, and 24001–65535 pixels, respectively. Data are represented as mean ± SEM. **F.** Viable cell numbers measured at D7 following daily treatment of Her2-dependent cells with vehicle, heparin low dose (lo; 5µg/ml), or heparin high dose (hi; 25µg/ml). n=4 replicates/group. ns=non-significant,

***$p<0.001$, ****$p<0.0001$. **G.** Forest plots of hazard ratios (HR) and 95% confidence intervals (CI) as a function of heparan sulfate/heparin or **H.** chondroitin sulfate/dermatan sulfate KEGG biosynthesis signatures in patients with breast cancer recurring within 10 years after initial treatment. Red dashed lines depict the shift in HR across 16 human datasets. Average HR highlighted in bold.

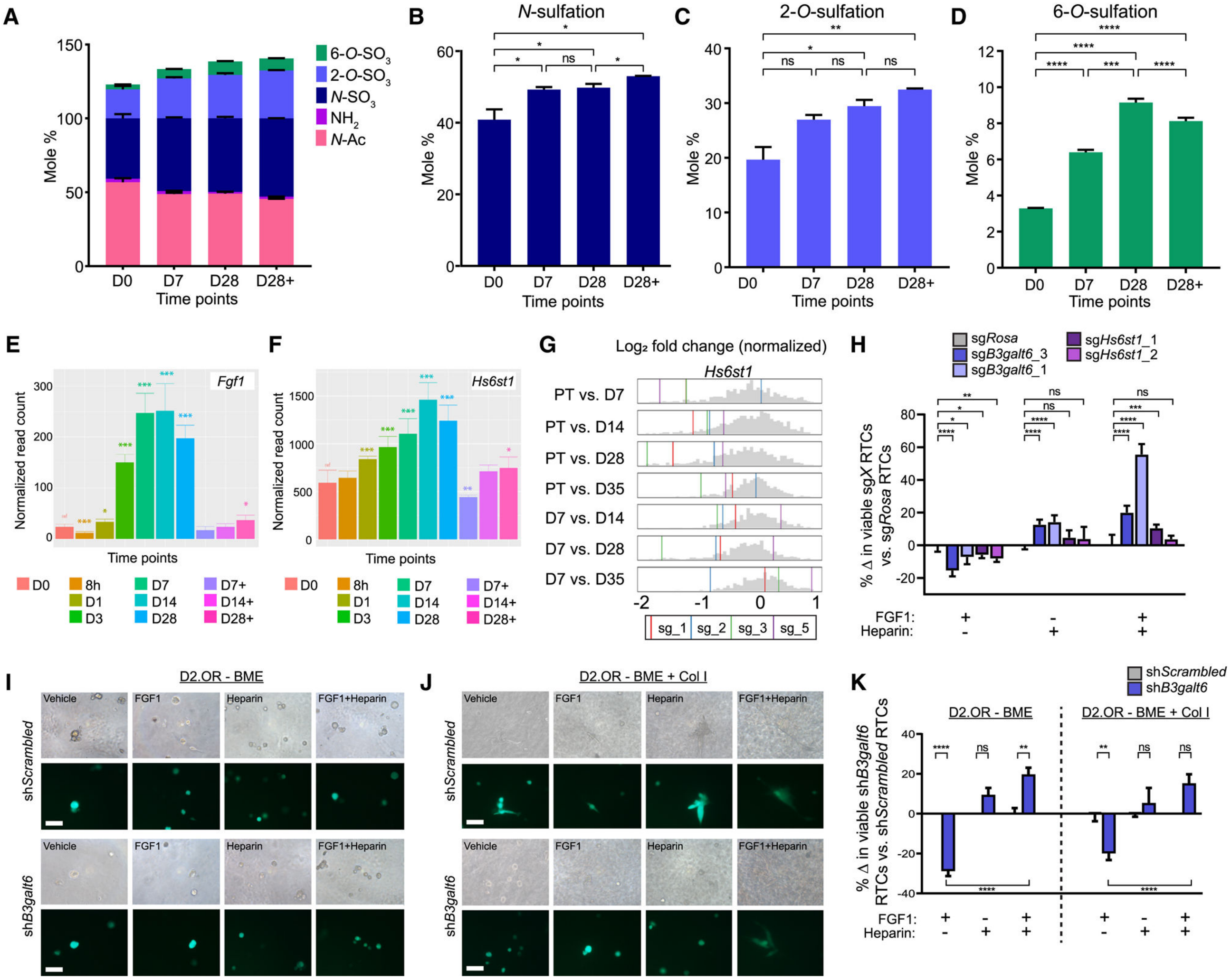


Fig. 7: Heparan sulfate 6-*O*-sulfation is selectively upregulated during dormancy and potentiates FGF1 signaling

A. LC/MS on D0 (*baseline*), D7, D28 (*deinduction*), and D28+ (*reinduction*) IVD samples identifying the molar percentages of modifications on disaccharide units comprising iduronic/glucuronic acid and glucosamine. Relative proportions of *N*-acetylglucosamine (*N*-Ac; light pink), unsubstituted glucosamine (NH_2; dark pink), **B.** *N*-sulfoglucosamine (*N*-S; dark blue), **C.** Uronyl-2-*O*-sulfates (2-*O*-SO_3; light blue), and **D.** Glucosaminyl-6-*O*-sulfates (6-*O*-SO_3; green) are depicted as mean ± SD. ns=non-significant, *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. **E.** Normalized read counts indicating reversible upregulation of *Fgf1* and **F.** *Hs6st1* at 8h, D1, D3, D7, D14, D28 (*deinduction*) time points. Asterisks indicate significant changes in normalized read counts vs. D0 (*baseline*) *$p<0.05$, **$p<0.01$,***$p<0.001$. **G.** Log_2 fold-change (normalized) gene-level effect size plots for *Hs6st1* in pairwise comparisons using PT or D7 as baselines. Grey histogram depicts the background distribution of sgRNAs in the CRISPR-Cas9 screen, colored vertical lines depict each sgRNA targeting *Hs6st1*. **H.** Viable RTC counts at D7 quantifying the percentage change relative to sg*Rosa* RTCs in sg*B3galt6*_3, sg*B3galt6*_1, sg*Hs6st1*_2, and sg*Hs6st1*_1 Her2-dependent-Cas9 cells treated with FGF1 (25ng/ml), heparin (25μg/

ml), or FGF1 (25ng/ml) + heparin (25µg/ml). n=6 replicates/group. ****$p<0.0001$. **I.** Brightfield and fluorescence images of sh*Scrambled* controls and sh*B3galt6* D2.OR cells following treatment with vehicle, FGF1 (25ng/ml), heparin (25µg/ml), or FGF1 (25ng/ml) + heparin (25µg/ml) in basement membrane extract (BME) or **J.** BME + collagen I (Col I). Scale bar=100µm. **K.** Viable sh*B3galt6* cell numbers normalized to sh*Scrambled* controls measured at D10 following treatment of D2.OR cells with FGF1 (25ng/ml), heparin (25µg/ml), or FGF1 (25ng/ml) + heparin (25µg/ml) in BME or BME + Col I. n=4 replicates/group. **$p<0.01$,****$p<0.0001$.

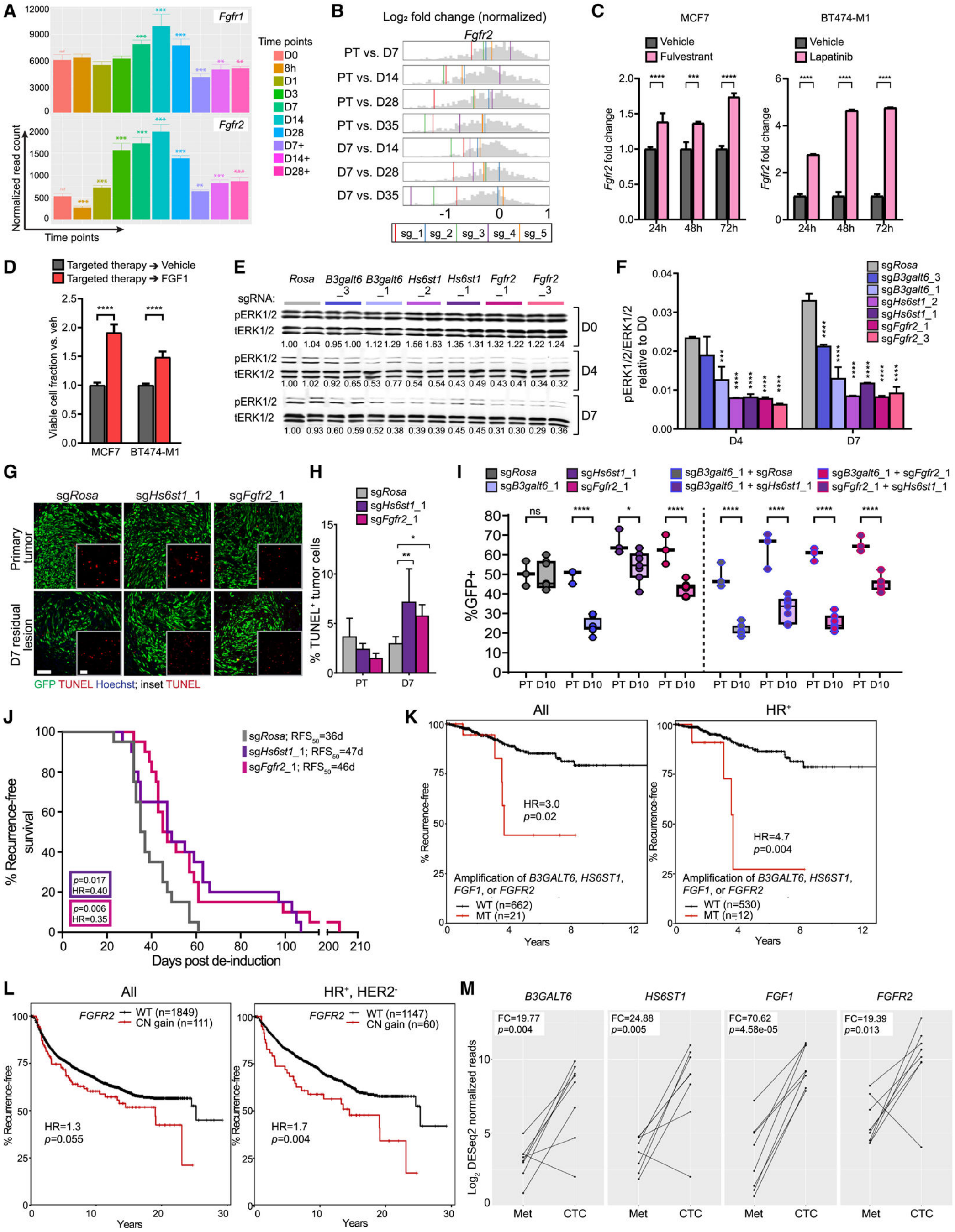


Fig. 8: Upregulation of heparan sulfate 6-*O*-sulfation during dormancy promotes RTC survival and recurrence in part by promoting FGFR2 signaling

A. Normalized read counts indicating the expression of *Fgfr1* and *Fgfr2*. Asterisks indicate significant changes in normalized read counts vs. D0 (*baseline*) **p*<0.05, ***p*<0.01, ****p*<0.001. **B.** Log$_2$ fold-change (normalized) gene-level effect size plots for *Fgfr2* in pairwise comparisons using PT or D7 as baselines. Grey histogram depicts the background distribution of sgRNAs in the CRISPR-Cas9 screen, colored vertical lines depict each sgRNA targeting *Fgfr2*. **C.** qRT-PCR for *FGFR2* transcripts in MCF7 and BT-474-M1 cells treated with vehicle (grey) or targeted therapies (pink), *i.e.*, Fulvestrant (100nM), and Lapatinib (400nM), respectively. Fold-change calculated relative to the vehicle controls. n=3 biological replicates/group. **D.** Viable cell numbers in 72h targeted therapy-treated cells measured after an additional 72h daily treatment with vehicle or hFGF1 (25ng/ml, orange) (normalized to vehicle-treated controls) **E.** Western blot analysis of phospho-ERK1/2 (pERK1/2) and total ERK1/2 (tERK1/2) levels in sg*Rosa*, sg*B3galt6*_3, sg*B3galt6*_1, sg*Hs6st1*_2, sg*Hs6st1*_1, sg*Fgfr2*_2, and sg*Fgfr2*_3 Her2-dependent-Cas9 cells at D0 (*baseline*), D4, and D7 (*deinduction*) time points. Numbers indicate relative quantification of pERK1/2/tERK1/2 signal normalized to sg*Rosa* levels within each time point. **F.** pERK1/2/tERK1/2 signal normalized to each sgRNA's baseline level at D0. Asterisks indicate significant changes vs. sg*Rosa* signal within the time point (*baseline*). Inset displays relative pERK1/2/tERK1/2 levels at D7 on a linear y-axis. *$p<0.05$, **$p<0.01$, ***$p<0.001$. **G.** Immunofluorescence and **H.** quantification for TUNEL (red) in Her2-dependent-Cas9 sg*Rosa*, sg*Hs6st1*_1, and sg*Fgfr2*_1 (green) PTs and D7 RLs. n=6 PT, 14 D7 RL/group. Scale bar=100μm. Quantification in the sg*Rosa* (grey), sg*Hs6st1*_1 (purple), and sg*Fgfr2*_1 (pink) groups is represented as mean ± SD. *$p<0.05$, **$p<0.01$. **I.** ddPCR data quantifying GFP+ cell percentage in sg*Rosa* (grey), sg*B3galt6*_1 (blue), sg*Hs6st1*_1 (purple), and sg*Fgfr2*_1 (pink) groups and sg*B3galt6*_1+sg*Rosa* (grey - blue outline), sg*B3galt6*_1+sg*Hs6st1*_1 (purple - blue outline), sg*B3galt6*_1+sg*Fgfr2* (pink - blue outline), and sg*Fgfr2*+sg*Hs6st1*_1 (purple - pink outline) combinatorial groups. Data are represented as median ± range. n=3 PT, 7 D10 RL/group. ns=non-significant, *$p<0.05$, ****$p<0.0001$. **J.** Kaplan-Meier analysis of recurrence-free survival for the sg*Rosa* (grey), sg*Hs6st1*_1 (purple), and sg*Fgfr2*_1 (pink) groups. n=20 mice/group. RFS_{50}=median time-to-recurrence. **K.** Kaplan-Meier analysis of recurrence-free survival in patients across breast cancer subtypes (left) and in the hormone receptor (HR)+ breast cancer subtype (right) with amplification of *B3GALT6*, *HS6ST1*, *FGF1*, or *FGFR2* in the TCGA-BRCA dataset. **L.** Kaplan-Meier analysis of recurrence-free survival in patients across breast cancer subtypes (left) and in the hormone receptor (HR)+, HER2− breast cancer subtype (right) with copy number gain of *FGFR2* in the METABRIC dataset. **M.** Log_2 DESeq normalized reads for *B3GALT6*, *HS6ST1*, *FGF1*, and *FGFR2* in paired CTC and metastasis samples from n=8 patients with breast cancer. FC=fold change.

KEY RESOURCES TABLE

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
Antibodies		
Rat anti-Ki67	Thermo Fisher Scientific	Cat#14–5698; RRID:AB_2865120
Rabbit anti-Her2	Cell Signaling	Cat#2165; RRID:AB_10692490
Rabbit anti-cleaved Caspase-3	Cell Signaling	Cat#9664; RRID:AB_2070042
Mouse anti-Chondroitin sulfate	Sigma	Cat#SAB4200696
Mouse anti-Heparan sulfate	Amsbio	Cat#370225
Rabbit anti-GFP	Cell Signaling	Cat#2956; RRID:AB_1196615
Mouse anti-GFP	Takara, Living Colors	Cat#632381; RRID:AB_2313808
Rabbit anti-phospho-ERK1/2 (Thr202/Tyr204)	Cell Signaling	Cat#9101; RRID:AB_331646
Rabbit anti-ERK1/2	Cell Signaling	Cat#9102; RRID:AB_330744
Mouse anti-b-Tubulin	BioGenex	Cat#MU122; RRID:AB_2335622
Goat anti-mouse IgG2a Alexa488	Thermo Fisher Scientific	Cat#A21131; RRID:AB_2535771
Goat anti-rabbit IgG Alexa488	Thermo Fisher Scientific	Cat#A11034; RRID:AB_2576217
Goat anti-rabbit IgG Alexa594	Thermo Fisher Scientific	Cat#A11012; RRID:AB_2534079
Goat anti-rat IgG Alexa568	Thermo Fisher Scientific	Cat#A11077; RRID:AB_2534121
Goat anti-mouse IgM Alexa568	Abcam	Cat#ab175702
anti-mouse 680LT	LI-COR Biosciences	Cat#925–68020; RRID:AB_2687826
anti-rabbit 800CW	LI-COR Biosciences	Cat#925–32211; RRID:AB_2651127
Bacterial and Virus Strains		
One Shot™ Stbl3™ Chemically Competent	Thermo Fisher Scientific	Cat#C737303
Lentivirus	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
mFGF1	R&D Systems	Cat#AF4686
Heparin	Fisher Scientific	Cat#AAA1619803
Fulvestrant	Selleck Chemicals	Cat#S1191
Lapatinib	Selleck Chemicals	Cat#S2111
Heparin lyase	Amsbio	Cat#AMS.HEP-ENZ III-S
Chondroitinase ABC	Amsbio	Cat#AMS.E1028–02
Critical Commercial Assays		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cell Titer 96 Non-Radioactive Cell Proliferation Assay	Promega	Cat#G4000
In Vivo EdU Click Kit 647	Sigma Aldrich	Cat#BCK647-IV-IM-S
Click-iT™ Plus TUNEL Assay for In Situ Apoptosis Detection, Alexa Fluor™ 647 dye	Thermo Fisher Scientific	Cat#C10640
Deposited Data		
MTB/TAN in vitro dormancy signature: RNA-sequencing data	This paper	Gene expression omnibus (GEO) accession no. GSE241252
MTB/TAN in vivo dormancy signature	Lewis A. Chodosh[7]	N/A
D2.OR 3D signature	Laila Ritsma[34]	N/A
Neoadjuvant chemotherapy response	Clemens B. Tempfer[21]; J Michael Dixon[22]; Andrew H Sims[23]; John Michael Dixon[24]; Jo Anne Zujewski[25]; Fatima Mechta-Grigoriou[26]; Lajos Pusztai[27]	N/A
Copy number data - METABRIC, TCGA-BRCA	cBioPortal[62, 63]	N/A
Circulating tumor cell data	Julie E. Lang[48]	Gene expression omnibus (GEO) accession no. GSE113890
Experimental Models: Cell Lines		
MTB/TAN-derived primary tumor cells	*MMTV-rtTA;TetO-HER2/neu* (*MTB/TAN*) transgenic mouse primary tumor; generated by our laboratory[59]	N/A
D2.OR	Mikala Egeblad	N/A
D2A1	Mikala Egeblad	N/A
MCF7	ATCC	Cat#HTB-22
BT474-M1	Mien-Chie Hung	N/A
Experimental Models: Organisms/Strains		
Mouse: *MTB/TAN*: *MMTV-rtTA;TetO-HER2/neu*	This laboratory[59]	N/A
Oligonucleotides		
List of sgRNAs used in the CRISPR/Cas9 screen	Table S1	N/A
ICE analysis primer sequence: sg*Rosa*	Sigma	F-GCGGGAGAAATGGATATGAA; R-GCACTTGCTCTCCCAAAGTC
ICE analysis primer sequence: sg*B3galt6*_3, sg*B3galt6*_1	Sigma	F-CAGGTCCGCAGAAAGGACAT; R-ACCACTCTGTTGTACCTGGC
ICE analysis primer sequence: sg*Hs6st*_2, sg*Hs6st*_1	Sigma	F-ATCCTTTACCAGTACGCGGG; R-AGTGACCCAGGAGGAAGTCT
ICE analysis primer sequence: sg*Fgfr2*_1	Sigma	F-CCTACTTGGGATTGCCAGCA; R-GGTTTCTGCAATCTGGACGC
ICE analysis primer sequence: sg*Fgfr2*_3	Sigma	F-GAAATGGGCCCACTGAGTCA; R-TTCGTGTCTCTCGGTTGTGG
Recombinant DNA		
LentiV_Cas9_puro	Christopher R Vakoc[60]	Addgene, Cat#108100; RRID:Addgene_108100

REAGENT or RESOURCE	SOURCE	IDENTIFIER
LRG2.1	Christopher R Vakoc[60]	Addgene, Cat#108098; RRID:Addgene_108098
LRG	Christopher R Vakoc[60]	Addgene, Cat#65656; RRID:Addgene_65656
LRmCherry2.1	Christopher R Vakoc[60]	Addgene, Cat#108099; RRID:Addgene_108099
pMD2.G	N/A	Addgene, Cat#12259; RRID:Addgene_12259
psPAX2	N/A	Addgene, Cat#12260; RRID:Addgene_12260
Software and Algorithms		
QuPath-0.3.0	N/A	https://qupath.github.io/