Bone vascular niche E-selectin induces mesenchymal–epithelial transition and Wnt activation in cancer cells to promote bone metastasis

Mark Esposito1, Nandini Mondal2,3,8, Todd M. Greco1,8, Yong Wei1, Chiara Spadazzi4, Song-Chang Lin5, Hanqiu Zheng1, Corey Cheung1, John L. Magnani6, Sue-Hwa Lin5, Ileana M. Cristea1, Robert Sackstein2,3,7 and Yibin Kang8,*

How disseminated tumour cells engage specific stromal components in distant organs for survival and outgrowth is a critical but poorly understood step of the metastatic cascade. Previous studies have demonstrated the importance of the epithelial–mesenchymal transition in promoting the cancer stem cell properties needed for metastasis initiation, whereas the reverse process of mesenchymal–epithelial transition is required for metastatic outgrowth. Here we report that this paradoxical requirement for the simultaneous induction of both mesenchymal–epithelial transition and cancer stem cell traits in disseminated tumour cells is provided by bone vascular niche E-selectin, whose direct binding to cancer cells promotes bone metastasis by inducing mesenchymal–epithelial transition and activating Wnt signalling. E-selectin binding activity mediated by the α1–3 fucosyltransferases Fut3/Fut6 and Glg1 are instrumental to the formation of bone metastasis. These findings provide unique insights into the functional role of E-selectin as a component of the vascular niche critical for metastatic colonization in bone.

A critical gap in our understanding of cancer metastasis lies between the initial dissemination of circulating tumour cells to a secondary organ and their successful outgrowth to metastatic lesions3–5. Two molecular themes central to metastasis initiation have emerged from recent studies: cellular plasticity mediated by epithelial—mesenchymal and mesenchymal—epithelial transitions (EMT and MET, respectively)6,7, and tumour–stromal niche interactions, which activate the properties of tumour-initiating or cancer stem cells through Wnt, Notch and other pathways8–10. Despite numerous studies, we do not yet understand how these programs interact to initiate metastasis.

The acquisition of mesenchymal traits through EMT has been shown to promote many pro-metastatic properties6–8. However, secondary metastatic tumours often match the epithelial state of the primary tumour10 and studies have suggested a requirement for MET during metastatic colonization11–14. Thus, it has been proposed that sequential EMT–MET are needed for tumour cells to colonize a distant organ1. Yet this hypothesis does not explain the paradox of how MET induction and stem-cell identity coexist when essential stem- and metastasis-associated factors such as Sox2 are suppressed in an epithelial state15,16. Furthermore, the causes of these sequential epithelial–mesenchymal–epithelial transitions remain only half described; the inducers of EMT are well-studied1 yet few inducers of MET are known.

For metastatic cells to colonize an organ, the inducers of MET must act in a context-dependent manner, which suggests that tumour–stromal interactions are instructive to MET. Clues from previous studies suggest that the bone vascular niche that hosts haematopoietic stem/progenitor cells (HSCs/HPCs) may contain such stromal signals, as this niche is co-opted by cancer cells during bone metastasis17,18. One factor essential to the function of the HSC/HPC vascular niche is endothelial selectin (E-selectin, SelE or CD62E). As originally defined, E-selectin functions to capture circulating leukocytes or tumour cells onto vascular endothelium under haemodynamic shear flow19, yet leukocyte recruitment is not altered by the ablation of E-selectin20. Further studies have shown that E-selectin plays a critical role in the creation of the bone marrow vascular niche that drives haematopoiesis and engagement of E-selectin with its ligands on haematopoietic progenitor cells21 leads to their proliferation22.

Although it is known that certain tumour cells bind to E-selectin in vitro23,24, genetic ablation of E-selectin has not been shown to affect metastasis in vivo; furthermore, our knowledge of the role of E-selectin receptor–ligand interactions in metastasis is incomplete. Here we report that Golgi glycoprotein 1 (Glg1 or E-selectin ligand 1) and glycoprotein E-selectin ligands created by the α1–3 fucosyltransferases 3 or 6 (Fut3/6) play key roles in mediating metastasis to bone by binding E-selectin, which induces a MET followed by stemness-enhancing Wnt signalling.

1Department of Molecular Biology, Princeton University, Princeton, NJ, USA. 2Department of Dermatology and Harvard Skin Disease Research Center, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA. 3Program of Excellence in Glycosciences, Harvard Medical School, Boston, MA, USA. 4Osteoncology and Rare Tumors Center, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori IRCCS, Meldola, Italy. 5Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. 6Glycomimetics Inc., Gaithersburg, MD, USA. 7Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA. 8These authors contributed equally; Nandini Mondal, Todd M. Greco. *e-mail: ykang@princeton.edu
Results
Enriched expression of E-selectin in bone specifically promotes bone metastasis. We first tested the correlation between in vitro E-selectin binding and metastatic propensity in vivo. To accurately quantify E-selectin binding across different cell lines, we designed an internally controlled flow cytometry assay (Supplementary Fig. 1a). The fidelity of this assay to detect E-selectin binding was verified using isotype-control IgG, a glycomimetic inhibitor of E-selectin (GMI-1271, uproleselan) and EDTA (Supplementary Fig. 1b). By applying this assay to a panel of isogenic MDA-MB-231 sub-lines with different organotropic metastatic abilities, we found that sublines with high bone or lung metastatic abilities generally have increased E-selectin binding ability (Supplementary Fig. 1c,d).

Whereas mouse and xenograft models of colon cancer or melanoma tested in E-selectin knockout (Sele<sup>−/−</sup>) mice have demonstrated that E-selectin is not essential for lung metastasis or primary-tumour formation, the effect of genetic ablation of E-selectin on bone metastasis has not been evaluated. We therefore crossed Sele<sup>−/−</sup> mice into the non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mouse background for xenograft studies of bone metastasis. Immunofluorescent staining of E-selectin and the endothelial marker CD31 revealed co-localization in the trabecular and cortical bone marrow space in wild-type (WT) mice (Fig. 1a), which is consistent with previous studies, and E-selectin staining was absent in the bone vasculature of Sele<sup>−/−</sup> NOD/SCID mice (Fig. 1b). Immunostaining and quantitative PCR (qPCR) demonstrated that E-selectin expression is significantly lower in the lung and liver vasculature compared with bone in both normal (Fig. 1b and Supplementary Fig. 1e) and inflammatory conditions (Supplementary Fig. 1e) known to induce E-selectin expression. Finally, qPCR analysis of pooled RNA collected from human tissues showed a two-fold greater expression of E-selectin in bone compared to lung or liver.
of CD31-normalized Sele messenger RNA in bones compared with lungs (Supplementary Fig. 1f).

We next analysed bone metastasis of the bone-tropic BM2 (cell line 1833) subline of MDA-MB-231 in E-selectin knockout or WT NOD/SCID mice. Bioluminescent imaging (BLI), X-ray and micro-computed tomography (μCT) analyses revealed that genetic knockout of E-selectin significantly attenuated the bone metastatic tumour burden (Fig. 1c–e). A similar result was observed when an in vivo–derived subline of the SUM159 breast cancer line that shows a strong bone metastatic propensity, named SUM159-M1a, was injected into Sle−/− mice compared with WT mice (Supplementary Fig. 1g–j). The primary tumour growth of SUM159-M1a was largely unaffected by E-selectin knockout, although an increase in tumour volume in Sle−/− mice emerged on the final day of measurement (Supplementary Fig. 1k), which is consistent with a previous report 1e.

As the lung metastatic LM2 subline of MDA-MB-231 has substantial E-selectin binding ability (Supplementary Fig. 1c), we used this cell line to test the effect of E-selectin knockout on lung metastasis. Knockout of E-selectin did not attenuate lung metastasis; rather, we observed a trend towards increasing total lung metastasis burden in Sle−/− NOD/SCID mice (Supplementary Fig. 1–n). We observed a similar result when the same lung metastasis experiment was performed using SUM159-M1a cells (Supplementary Fig. 1o,p). Together, these data establish that E-selectin is expressed at an elevated level in the bone vasculature and is specifically important for the development of bone metastasis.

Both α1–3 Fut3 and Fut6 promote E-selectin-mediated bone metastasis. Binding of E-selectin to a cell requires the presence of sialyl Lewis X or A (sLeX/A) tetrasaccharides at the termini of the cell-surface glycolipids, glycoprotein O-glycans (Ser/Thr-linked) or N-glycans (Asn-linked) 31. Six fucosyltransferases (Fut3, -4, -5, -6, -7 and -9) are capable of performing the final α1–3 or α1–4 fucosylation to generate sLex (refs. 32–36). Importantly, Fut4, -7 and -9 are present in both the human and murine genomes, whereas Fut3, -5 and -6 are present only in humans 3e. The contributions of these enzymes to tumorigenesis and organotropic metastasis have not been systematically analysed. Ectopic expression of each Fut in the MDA-MB-231 cell line did not affect the proliferation rates (Supplementary Fig. 2a,b), whereas flow cytometry showed that Fut3, -5, -6 and -7 increase E-selectin binding by several orders of magnitude while Fut4 and Fut9 increase binding by two- to fivefold (Fig. 2a). A similar binding pattern was observed in the SUM159-M1a cell line following the ectopic expression of Fut enzymes (Supplementary Fig. 2c,d).

Intracardiac injection of M1a–Fut cells demonstrated that only the human-specific Fut enzymes (Fut3, -5 and -6) and not Fut7 promoted bone metastasis (Fig. 2b,c and Supplementary Fig. 2d,f), which is supported by previous findings that Fut6 mediates increased engraftment of human mesenchymal stem cells (MSCs) and human prostate cancer cells in bone 3h,3i. Clinical analysis of The Cancer Genome Atlas (TCGA) data next revealed that Fut5 is expressed at low levels in both normal and cancerous human breast tissues compared with more abundant expression of other Fut enzymes (Supplementary Fig. 2g).

To analyse which Fut enzymes endogenously contribute to increased E-selectin binding, the MDA-MB-231 cell line was fluorescence-activated cell sorted into the top and bottom decile of E-selectin binding (Supplementary Fig. 2h). After five passages, E-selectin binding levels were maintained at respectively higher or lower levels (Supplementary Fig. 2i) and qPCR revealed that Fut3 and Fut6 expression positively correlated with E-selectin binding, whereas Fut4 was negatively correlated (Fig. 2d). A similar pattern was observed between different sublines of MDA-MB-231 cells, with elevated Fut6 and lower Fut4 expression in a strongly bone metastatic subline (SCP25; Supplementary Fig. 2j).

We next tested whether the bone metastasis-promoting effects of Fut3 and -6 replicated in a spontaneous multi-organ metastasis model. SUM159-M1a cells expressing either of these bone metastasis-promoting Fut enzymes or Fut4 and -7 as negative controls were implanted in the mammary fat pad of NOD/SCID gamma mice. Ectopic Fut3, -4, -6 and -7 expression did not affect primary tumour growth (Fig. 2e), whereas only Fut3 and Fut6 promoted spontaneous bone metastasis and none of the enzymes affected spontaneous lung metastasis (Fig. 2f–h and Supplementary Table 1). To test whether these bone metastasis effects depend on Fut catalytic activity, we generated three catalytic mutants of Fut3 (Fig. 3a). E-selectin binding analysis confirmed that each mutant was catalytically inactive (Fig. 3b). In contrast to the increased bone metastasis caused by overexpression of Fut3, the mutant forms of Fut5 did not promote bone metastasis formation (Fig. 3c–f). In particular, the Y315–stop mutant reduced bone metastasis based on BLI measurements, possibly due to dominant negative effects on WT Fut3 and Fut6.

N-glycoprotein capture mass spectrometry identifies candidate substrates of human Fut enzymes. Given the data showing that increased expression of Fut3 or Fut6 promotes bone metastasis, non-fucosylated glycan substrates of these enzymes must exist in both cell lines (Supplementary Fig. 3a, top). We determined that these are displayed on N-glycans, as M1a and BM2 cells grown with either tunicamycin or 1-deoxymannojirimycin, which block N-linked glycosylation, demonstrated reduced E-selectin binding, whereas D,L-threo-PDMP, an inhibitor of glycosylceramide synthase, moderately reduced E-selectin binding to BM2 and did not affect E-selectin binding to M1a cells (Supplementary Fig. 3b,c). Furthermore, GMI-1271, a sLeX mimic, was capable of blocking E-selectin binding to M1a cells expressing each Fut enzyme (Supplementary Fig. 3d), indicating these were conventional E-selectin ligands.

We therefore adapted a solid-phase N-glycoprotein enrichment protocol to comprehensively profile cell surface glycoproteins and identify candidate Fut enzyme substrates (Supplementary Fig. 3a, bottom) 3j,3l. Imaging and flow cytometry controls demonstrated effective ligation of extracellular oxidized glycoproteins (Supplementary Fig. 4a–c). N-glycoproteins were then isolated from MDA-MB-231 and SUM159-M1a cells through: (1) trypsin digestion to remove non-glycosylated peptides, (2) N-glycopeptide release with PNGase F, (3) liquid chromatography–tandem mass spectrometry analysis and (4) computational filtering if peptides that contained one asparagine that was deamidated from PNGase F digestion matched the known N-X/S/T glycosylation motif and were detected in either all four samples or in both samples from a single cell line. Filtering by these criteria resulted in the identification of 1,037 unique N-glycosites that mapped to 541 proteins (Fig. 4a and Supplementary Table 2).

The 361 N-glycosylated proteins shared between the MDA-MB-231 and M1a cells were cross-referenced with reported E-selectin ligands and ranked according to individual N-glycopeptide detection across all four runs by their respective normalized intensities (Supplementary Table 3). The top candidate identified was CD44, which when decorated with sLeX is known as haematopoietic cell E-/L-selectin ligand (HCELL). HCELL has been attributed as the major E-selectin binding ligand on metastatic cells 3m,3n,3o,3p and HPCs 3l. To determine whether HCELL was responsible for E-selectin binding, we utilized lentiviral clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR associated protein 9 (Cas9) system to knock out CD44 in the BM2 cell line (Fig. 4b). Loss of CD44 expression did not decrease E-selectin binding in vitro (Fig. 4c) or inhibit bone metastasis in vivo (Supplementary Fig. 4d,e).
Fig. 2 | Specific α1–3 fucosyltransferases (Fut3 and Fut6) promote bone metastasis. a, Comparative flow cytometry analysis of E-selectin binding to MDA-MB-231 cells with stable ectopic expression of α1-3 Fut enzymes. MDA-MB-231-red fluorescent protein (RFP) was used as an internal control. The data are representative of four independent experiments. b, BLI quantification of the bone metastasis burden following intracardiac injection of M1a cells stably expressing each Fut enzyme into Nu/Nu mice. Statistics were performed using two-sided Mann–Whitney U tests at Day 21; n = 6 (Fut3, Fut4 and Fut6), 5 (Fut5 and Fut9), 3 (Fut7) and 10 (Vector) mice. c, Representative BLI (left and middle) and X-ray (right) images of the bone lesions from b. The white arrows indicate osteolytic lesions. d, Analysis by qPCR of the endogenous α1–3 Fut mRNA levels in the parental and sorted MDA-MB-231 cells with differential E-selectin binding abilities. Fut5, -7 and -9 were not detectable (N.D.) in any of the cell lines. n = 3 technical replicates. e, Tumour volume measurements after orthotopic injection of M1a cells stably expressing each relevant Fut enzyme into NOD/SCID gamma mice. Two-sided Student’s t-test at Day 39; n.s, not significant; n = 6 mice per group. f-g, BLI quantification of the spontaneous bone (f) and lung (g) metastasis burden. The plot elements include the median, box for the interquartile range and spikes to the upper and lower adjacent values. One-sided Mann–Whitney U tests; n = 11 lungs and 22 hindlimbs (Fut3 and Fut 4), 6 lungs and 12 hindlimbs (Fut6 and Fut7), and 12 lungs and 24 hindlimbs (Vector). h, Representative BLI images of bone and lung tissues from each experimental group. e–h, The experiment was performed once. Data represent the mean ± s.e.m. No statistically significant differences (P > 0.05) were observed between the Fut4, Fut7 and Fut9 groups versus the Vector group in b, the Fut4 and Fut7 groups versus the Vector group in f, all group comparisons in e and all of the groups versus the Vector group in g.
Fig. 3 | In vitro and in vivo characterization of Fut3 catalytic mutants. a. Western blot detection of three Fut3 catalytic mutants compared with WT Fut3 ectopically expressed in SUM159-M1a cells. b. Comparative flow cytometry analysis of E-selectin binding to M1a cells expressing each Fut3 mutant using SUM159-RFP as an internal control. The data in a b are representative of three independent experiments. c. BLI quantification of the bone metastasis burden following intracardiac injection of M1a cells stably expressing each Fut3 mutant compared with vector and WT Fut3. Two-sided Mann–Whitney U tests; n = 8 (Fut3 and Y315-stop) and 9 (E247K and Vector) mice per group. d. Representative BLI (left and middle) and X-ray (right) images from c. e, f. Quantification of the osteolytic area (e) and the number of osteolytic lesions (f) in the hindlimbs of Nu/Nu mice receiving an intracardiac injection of M1a cells expressing each Fut3 mutant. Two-sided Mann–Whitney U tests; n = 16 (Fut3 and Y315-stop) and 18 (E247K and Vector) hindlimbs per group. e–f. The experiment was performed once. Data represent the mean ± s.e.m. Unprocessed original scans of the blots in a are shown Supplementary Fig. 9.
Glg1 is functionally important to E-selectin binding of bone metastatic breast cancer cells. N-glycosylation of the second candidate Glg1 (also known as E-selectin ligand 1, Es1; Fig. 4d,e) was detected at two different residues, with Asn165 detected across all of the samples (Supplementary Table 3). The site and peptide was confirmed by tandem mass spectrometry fragmentation and (as well as) similar elution profiles and extracted ion chromatograms from both cell lines (Fig. 4d and Supplementary Fig. 4f). Previous research has identified a role for Glg1 in HPCs by maintaining homeostasis or proliferation, possibly in an E-selectin-dependent manner. Although Glg1 has been detected in metastatic prostate cells, the functional importance of Glg1 in cancer has not been investigated.

To test whether Glg1 expression affected E-selectin binding, isoform-specific primers for the three main Glg1 variants (Fig. 4e) were used to identify a role for Glg1 in HPCs by maintaining homeostasis or proliferation, possibly in an E-selectin-dependent manner. Although Glg1 has been detected in metastatic prostate cells, the functional importance of Glg1 in cancer has not been investigated.
To assess whether Glg1 directly binds E-selectin, we performed a representative of three independent experiments. Scale bars, 5 μm.

**Fig. 5 | E-selectin ligands co-localize with Glg1 at the cell surface.**

**a.** Total internal reflectance fluorescence microscopy at the critical angle using a spinning disc was performed on SUM159-M1a cells probed with recombinant E-selectin–Fc and anti-Glg1. **b.** Confocal z-slice of M1a cells probed with E-selectin–Fc, anti-Glg1 and Hoechst. **a, b.** The data are representative of three independent experiments. Scale bars, 5μm.

![Image](https://example.com/image.png)

They were used for qPCR. We found that variants 1 and 3 were enriched in E-selectin high-binding cells (E-high and SCP25), whereas variant 2, which harbours an internal deletion located eight amino acids upstream from the putative Asn165 modification site, was inversely correlated to the E-selectin binding ability (Supplementary Fig. 4g,h). We ectopically expressed Glg1 variants 1 and 3 in SUM159, M1a, MDA-MB-231 and BM2 cells, and knocked out Glg1 in M1a and BM2 cells using lentiviral CRISPR–Cas9 (Fig. 4l,g and Supplementary Fig. 4i–k). The comparative flow cytometry assay revealed that ectopic expression of Glg1 variant 1 and variant 3 increased E-selectin binding to varying degrees in different cell lines, whereas population-level CRISPR–Cas9-mediated knockout of Glg1 reduced E-selectin binding by approximately 50% in the M1a cells and about 20% in the BM2 cell line (Fig. 4h and Supplementary Fig. 4i). Consistent with the result using BM2 cells (Fig. 4c), we observed that CD44 knockout also did not affect E-selectin binding in the M1a cell line (Fig. 4h).

To serve as a bona fide E-selectin ligand, Glg1 must be expressed on the cell membrane and be capable of binding E-selectin. Total internal reflectance fluorescence imaging was performed at the critical angle on M1a cells probed with anti-Glg1 antibody and E-selectin–IgG demonstrated that the Glg1 and E-selectin ligands are displayed on the cell surface, co-localized to the same puncta (Fig. 5a). Furthermore, western blotting of Glg1 revealed its presence in membrane fractions of M1a cells (Supplementary Fig. 5a). Confocal microscopy of permeabilized M1a cells also demonstrated colocalization of Glg1 and E-selectin binding throughout the cell, both at the cell membrane and in internal structures (including the Golgi apparatus and endoplasmic reticulum; Fig. 5b). To assess whether Glg1 directly binds E-selectin, we performed a western blot under denaturing conditions on cell membrane fractions from M1a Glg1 knockout, Glg1 overexpressing, and Fut3–7 and Fut9 overexpressing cells using E-selectin–IgG as a probe with Fut6-transfected prostate cancer cells (PCR1) serving as a positive control. Although E-selectin–IgG-reactive bands were detectable in the cell membrane fractions of cells expressing Fut3, -5, -6 and -7, E-selectin–IgG-reactive bands were not detected in control, Glg1 knockout, Glg1 overexpression or Fut4 and Fut9 membrane fractions (Supplementary Fig. 5b) despite evidence of E-selectin binding in flow cytometry and immunofluorescence microscopy. Immunoprecipitation of M1a lysates by E-selectin–IgG followed by probing with Glg1 antibody further confirmed that an E-selectin interaction with Glg1 was not detectable using traditional biochemical methods (Supplementary Fig. 5c). Thus, although the data indicate that: (1) N-glycan decoration of proteins but not glycosphingolipids is critical to E-selectin binding, (2) Glg1 is expressed on the cell surface, is N-glycosylated and co-localizes with E-selectin binding, and (3) genetic manipulation of Glg1 leads to significant changes in E-selectin binding levels, we did not obtain evidence that Glg1 directly binds E-selectin.

We next tested whether Glg1 ablation affects metastasis in vivo. Two lentiviral short hairpin RNA (shRNA) constructs targeting all of the Glg1 splice variants reduced Glg1-expression levels in BM2 cells (Supplementary Fig. 6a). Bioluminescent imaging following intracardiac injection confirmed that Glg1 knockdown reduced the bone metastasis burden (Fig. 6a). Similarly, CRISPR–Cas9-mediated knockout of Glg1 in BM2 and M1a cells resulted in a significant reduction of bone metastasis based on BLI, X-ray and μCT analyses of bone lesions (Fig. 6b–d and Supplementary Fig. 6b–f). In contrast, Glg1 knockdown in LM2 cells (Supplementary Fig. 6g) did not affect the progression of lung metastasis or survival of the animals (Supplementary Fig. 6h,i). The results were in line with the same observations made with either E-selectin knockout (Fig. 1) or Fut overexpression (Fig. 2).

**Pharmaceutical targeting of E-selectin reduces bone metastasis.**

Given the identification of E-selectin as a pro-metastatic receptor of the bone vascular niche, the binding of which is predominantly governed by Fut3, Fut6 and Glg1, we sought to assess whether this relationship could predict bone metastasis in human breast cancer. Patients in the NKI-29547 and EMC-MSK datasets were stratified according to the median expression of Glg1 and Fut3 in these patients often manifest at multiple organ sites. We further assessed the prognostic potential of each candidate E-selectin ligand identified by mass spectrometry (Supplementary Table 3) in organ-specific metastasis and found the only significant relationship to be that of Glg1 as a poor prognosis indicator in bone metastasis (Fig. 6e and Supplementary Table 4h).

These results suggest that therapeutic inhibition of E-selectin in patients with tumours expressing a high level of Fut3 or Fut6 and Glg1 may slow down or prevent the progression of bone metastasis. To test this idea, mice were treated twice daily with GMI-1271 at 20 mg kg⁻¹ for 14 d immediately following intracardiac injection with BM2 cells. Bone metastasis-associated bone degradation was attenuated in mice treated with GMI-1271 (Fig. 6f–h). A distinct survival advantage was observed for the mice receiving GMI-1271 (Fig. 6i) and post-mortem μCT bone-volume scans revealed a greater preservation of bone tissue in these mice compared with mice treated with PBS (Fig. 6j).
E-selectin binding promotes a specific MET program. Although the majority of in vitro studies have speculated that E-selectin may promote metastasis by arresting tumour cells in the vasculature of target organs through adhesive interactions, this mechanism is unlikely to be important to bone metastasis as the blood flow in the sinusoidal endothelium of bone marrow has been noted to be particularly slow or even stationary. Instead, we considered the possibility that E-selectin engagement of metastatic cells to the endosteal endothelium during the initial phase of bone metastatic colonization (as visualized in Supplementary Fig. 7a) may induce molecular changes in cancer cells that promote bone colonization. This hypothesis was supported by the observation that the growing edge of BM2 bone lesions showed strong Glg1 expression and were associated with a high density of E-selectin vascularization (Supplementary Fig. 7b).

**Fig. 6** | Glg1 is required to support bone-metastasis progression. 

**a.** BLI quantification of the bone metastasis burden after intracardiac injection of BM2 cells stably expressing Glg1-targeting or control shRNA and untransduced (parental) BM2 cells into Nu/Nu mice. Two-sided Mann–Whitney U tests were used to compare the BLI signals of shGlg1-2 (purple) and shGlg1-3 (red) with the control; n = 8 mice per group. **b.** Representative BLI, X-ray and µCT images of mice from mice injected with BM2 cells modified by CRISPR–Cas9-mediated Glg1 knockout. The white arrows indicate the osteolytic bone lesions. **c.** The number of lesions per hindlimb (l) and average osteolytic area (d) were quantified from the X-ray images in b. Two-sided Mann–Whitney U tests; n = 17 hindlimbs per group. The data in a–d are representative of two independent experiments. **e.** Kaplan–Meier organ-specific metastasis-free survival curves of ER+ breast cancer patients in the EMC-MSK dataset stratified according to the median expression level of Glg1 mRNA. Two-sided Cox’s proportional hazards model; n = 244. **f.** Representative BLI, X-ray and µCT images of mice treated with either GMI-1271 or PBS control after intracardiac injection of BM2 cells to generate bone metastasis in Nu/Nu mice. Two-sided Mann–Whitney U tests; n = 9 mice per group. **g.** Number of lesions per hindlimb (l) and total osteolytic area (d) were quantified between treatments. Two-sided Mann–Whitney U tests; n = 12 hindlimbs per group. **h.** Mice from f were censored after becoming moribund and followed for Kaplan–Meier survival curve analysis. Two-sided Cox’s proportional hazards model; n = 9 mice per group. **i.** Bones from moribund mice collected on Day 39 (PBS, n = 3) or Day 44 (GMI-1271, n = 4) were analysed by µCT and the trabecular bone volume from 4 mm above and below the knee joint was quantified. Two-sided Mann–Whitney U test. f–j. The experiment was performed once. Data represent the mean ± s.e.m.; HR, hazard ratio.
To mimic the engagement of tumour cells to E-selectin, BM2 or M1a cells were seeded on E-selectin- or IgG-coated plates and subjected to spinning-disc microscopy over a period of 40 h. Dramatic differences in cell behaviour and morphology emerged within 3 h of seeding (Supplementary Video 1). Both BM2 and M1a cells migrated along the E-selectin-coated dishes forming three-dimensional clusters ranging in thickness from tens to hundreds of cells, whereas cells that were seeded on IgG plates formed...
simple monolayers as in normal tissue culture conditions (Fig. 7a). Immunofluorescence analysis of M1a spheres revealed enriched N-cadherin localization at cell–cell boundaries compared with diffuse staining in cells seeded over IgG. Furthermore, the tight junction protein-1 (TJP-1 or ZO-1) was only detectable in cells seeded on E-selectin (Fig. 7a). Staining for EpCam and Keratin-14 revealed an enhanced shift towards an epithelial state in BM2 spheres (Fig. 7b and Supplementary Fig. 7c).

Microarray and gene set enrichment analysis (GSEA) performed on cells that were seeded for 24 h on either E-selectin- or IgG-coated plates revealed an inverse enrichment of gene signatures related to EMT (Fig. 7c and Supplementary Fig. 7d,e). Unlike traditional models of EMT, E-selectin-induced MET did not affect the RNA expression of the master transcriptional regulators of EMT, such as Snai1/2, Twist1/2 and Zeb1/2 (Fig. 7d). Furthermore, staining of the EMT marker N-cadherin did not show reduced expression but rather altered the localization and apparent molecular weight (Fig. 7a and Supplementary Fig. 7f), whereas protein levels of the Slug transcription factor were much lower after binding E-selectin (Supplementary Fig. 7f). Together, these observations pointed towards a non-canonical MET program that is not the binary opposite of traditional EMT programs. Supporting this, extraction of the enrichment-score genes in the Sarrio-EMT and Hallmark EMT gene sets (Fig. 7c) showed that the majority of the upregulated EMT-associated genes were involved in immune-related processes, whereas the downregulated genes were largely secreted or extracellular proteins (Fig. 7c and Supplementary Table 5). These results indicate that E-selectin binding induces a non-canonical MET-like shift in cancer cells. Finally, staining for epithelial markers revealed ubiquitous E-cadherin and intermittent EpCam staining in BM2 bone lesions (Supplementary Fig. 7g,h), supporting the occurrence of MET during bone metastasis in vivo.

E-selectin-induced MET activates Wnt signalling. The three most downregulated mesenchymal-associated genes (Dkk1, Ctgf and Cyr61) after E-selectin-induced MET encode secreted Wnt repressors3 (Fig. 7f). We next investigated whether repression of these Wnt inhibitors by E-selectin-induced MET could activate Wnt signalling, as Wnt signalling has been reported to promote cancer stem cell activities during metastasis11. To this end, we introduced a 7×TCF–GFP Wnt reporter plasmid44 into BM2 cells (BM2-TGC). Cells seeded on E-selectin for 48 h activated this Wnt reporter to levels similar to cells seeded on IgG and treated with recombinant Wnt3a (Fig. 8a). Live imaging revealed that MET occurred in the first several hours, whereas Wnt activation occurred 30 h after binding (Supplementary Video 1). Quantitative PCR revealed that multiple Wnt target genes, including those related to cancer stem cells (for example, Sox2 and Sox9) were induced by plating tumour cells on E-selectin, whereas the canonical EMT markers Vim, Zeb1 and Zeb2 remained mostly unchanged (Fig. 8b), consistent with results from gene expression profiling analysis (Fig. 7d), thus confirming that E-selectin activated canonical Wnt signalling. Ex vivo microscopy of BM2-TGC bone lesions demonstrated that Wnt signalling was active in cells that were in a high density of CD31+ vasculature (Fig. 8c and Supplementary Fig. 8a).

Ectopic expression of Dkk1–FLAG in BM2-TGC cells (Supplementary Fig. 8b) blocked Wnt activation following exposure to Wnt3a-conditioned media but exerted a lesser effect on E-selectin-mediated Wnt signalling. In contrast, inhibitors of β-catenin–TCF binding (ICG-001 and LF3) disrupted E-selectin-mediated Wnt signalling while exerting less of an effect on paracrine-mediated Wnt activation (Supplementary Fig. 8c). Treatment of BM2 bone metastases with LF3 did not exhibit observable toxicity but reduced the number of osteolytic lesions (Supplementary Fig. 8d–f).

Recent glycometric analysis has found a strong correlation between extracellular fucosylation and epithelial features45. Quantitative PCR analysis of the levels of Glg1-variant 3, Fut3 and Fut6 48 h after seeding on E-selectin confirmed that E-selectin-induced MET led to increased expression of the genes responsible for the binding of tumour cells to E-selectin (Fig. 8d), indicating the existence of a positive feedback loop. This finding is consistent with the stronger Glg1 staining in metastatic cancer cells in close contact with the bone vasculature (Supplementary Fig. 7b). To investigate whether this Wnt activation contributed to cancer stem cell identity, we transduced BM2 cells with a Sox2/Oct4–mCherry (SORE6–mCherry) cancer stem cell reporter53. When plated on IgG, 11% of these cells were positive for Sox2/Oct4 but this increased to 31.7% after binding to E-selectin-coated plates, indicating an increase in cancer stemness activity following binding to E-selectin (Fig. 8e).

Finally, we wanted to understand whether this mechanism extended to other models of bone metastasis and whether the discovered mechanism is relevant to clinical bone metastasis patients. We therefore used in vivo selection to derive a bone metastatic subline (named ob1) from the DU145 prostate cancer cell line54. Flow cytometry revealed this subline binds E-selectin (Supplementary Fig. 8g). Similar to breast cancer studies, E-selectin knockout mice exhibited lower bone metastasis colonization and survived longer than WT hosts following intracardiac injection of DU145-ob1 (Supplementary Fig. 8h,i). Plating of these cells on E-selectin-coated plates revealed a similar phenomenon of clustering and a shift towards epithelial features (Supplementary Fig. 8j). Furthermore, immunostaining of E-selectin, Glg1 and Ki67 in bone metastasis biopsies from prostate cancer patients demonstrated an association between E-selectin expression and Glg1+ tumour cells that are proliferative (Ki67+; Supplementary Fig. 8k).

Discussion

Although numerous studies have established that EMT is often necessary for the escape from a primary tumour and others have shown that these cells must revert to an epithelial state to successfully colonize an organ11–13, the lack of evidence for how MET is induced, especially in the context-dependent manner required for metastatic colonization in a distant organ, has resulted in considerable controversy within the field. Here we provide evidence that a unique stromal cue—binding to E-selectin via Fut3/6 and Glg1 expressed by bone metastatic cells—induces MET to facilitate bone metastasis. The data further indicate that this non-canonical MET activates Wnt signalling to promote stemness via Sox2/9 expression and further increases Glg1 and Fut3/6 expression (Fig. 8f).

E-selectin-induced MET is a non-canonical program compared with MET inducers such as miR-200 that target the master transcription factors of EMT55. Rather, a key function of E-selectin-induced MET is the activation of Wnt signalling, a pathway linked to self-renewal, cancer stem cell traits and EMT induction56. This reported link between Wnt-induced stemness and EMT is incompatible with the requirement for MET during metastatic colonization, which therefore raises an interesting paradox: how can MET and cancer stemness coexist during metastatic colonization? By linking E-selectin-induced MET with Wnt signalling and Sox2/9 induction, we show how E-selectin engagement could resolve this paradox in bone metastasis, which is akin to recent findings showing that Prrx1 uncouples EMT from stemness traits during lung metastasis57.

Our discovery of the crucial role of E-selectin in bone metastasis raises the question of why no previous studies have shown a functional effect of E-selectin. Three discoveries made here explain a possible cause: (1) Fut3/6 are responsible for generating functional E-selectin ligands, thus implying that murine models of bone metastasis may not be dependent on E-selectin, (2) E-selectin-induced Wnt activation occurred 30 h after binding, whereas previous in vivo studies of E-selectin were performed within 6 h of injection58 and (3) the organ-specific expression pattern of E-selectin indicates
it is most relevant to bone metastasis as compared with lung or liver metastasis. Interestingly, in a transgenic mouse model of E-selectin under the control of the β-actin promoter, the murine B16F10 cell line could be redirected to the liver only if both Fut3 and E-selectin were ectopically expressed in the cell line and liver endothelium.

Our discovery that Glg1 is instrumental to MET during bone metastasis and co-localizes with E-selectin ligands opens numerous avenues for further research to elucidate whether it chaperones ligand maturation or expression, and how it may function as an intermediate between E-selectin binding and MET induction. The common dependence of MDA-MB-231, SUM159 breast cancer...
cells and DU145 prostate cancer cells on E-selectin binding to achieve bone metastasis, together with the prognostic significance of Glg1 and Fut3/Fut6 in clinical bone metastasis, indicate that the discovered pathway is probably a conserved mechanism that facilitates tumour–endothelial interaction during metastatic seeding in bone. Therapeutic agents targeting the E-selectin ligand-mediated pathway, such as LF3 or GM1-1271, may be developed for the treatment of bone metastasis.

**Online content**

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41556-019-0309-2.

Received: 14 January 2019; Accepted: 7 March 2019; Published online: 15 April 2019

**References**

1. Valetyan, S. & Weinberg, R. A. Tumor metastasis: molecular insights and evolving paradigms. Cell 147, 275–292 (2011).
2. Massague, J. & Obenauf, A. C. Metastatic colonization by circulating tumour cells. Nature 529, 298–306 (2016).
3. Yang, J. & Weinberg, R. A. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Dev. Cell 14, 818–829 (2008).
4. Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. EMT: 2016. Cell 166, 21–45 (2016).
5. Espósito, M., Guise, T. & Kang, Y. The biology of bone metastasis. Cold. Spring Harb. Perspect. Med. 8, a031252 (2018).
6. Oskarsson, T., Batlle, E. & Massague, J. Metastatic stem cells: sources, niches, and vital pathways. Cell Stem Cell 14, 306–321 (2014).
7. Yang, J. et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell 117, 927–939 (2004).
8. Mani, S. A. et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 133, 704–715 (2008).
9. Labelle, M., Begum, S. & Hynes, R. O. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. Cancer Cell 20, 576–590 (2011).
10. Brabletz, T. et al. Variable β-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. Proc. Natl Acad. Sci. USA 98, 10356–10361 (2001).
11. Ocana, O. H. et al. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. Cancer Cell 25, 10433–10444 (2010).
12. Ossola, R. et al. Biomarker validation in blood specimens by selected reaction monitoring mass spectrometry of N-glycans. Methods Mol. Biol. 728, 179–194 (2011).
13. Weeks, M. P. et al. Proteomic plasma membrane profiling reveals an essential role for gp96 in the cell surface expression of LDLR family members, including the LDL receptor and LRP6. J. Proteome Res. 11, 1475–1482 (2012).
14. Price, T. T. et al. Dormant breast cancer micrometastases reside in specific bone marrow niches that regulate their transit to and from bone. Sci. Transl. Med. 8, 340ra373 (2016).
15. Morita, Y. et al. E-selectin targeting PE-Glylated-thioapoptant prevents breast cancer metastases. Mol. Ther. Nucleic Acids 5, e399 (2016).
54. Fuerer, C. & Nusse, R. Lentiviral vectors to probe and manipulate the Wnt signaling pathway. PLoS One 5, e9370 (2010).
55. Holst, S. et al. N-glycosylation profiling of colorectal cancer cell lines reveals association of fucosylation with differentiation and caudal type homebox 1 (CDX1)/Villin mRNA expression. Mol. Cell. Proteomics 15, 124–140 (2016).
56. Stone, K. R., Mickey, D. D., Wunderli, H., Mickey, G. H. & Paulson, D. F. Isolation of a human prostate carcinoma cell line (DU 145). Int. J. Cancer 21, 274–281 (1978).
57. Korpal, M., Lee, E. S., Hu, G. & Kang, Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. J. Biol. Chem. 283, 14910–14914 (2008).
58. Anastas, J. N. & Moon, R. T. WNT signalling pathways as therapeutic targets in cancer. Nat. Rev. Cancer 13, 11–26 (2013).
59. Sipkins, D. A. et al. In vivo imaging of specialized bone marrow endothelial microdomains for tumour engraftment. Nature 435, 969–973 (2005).
60. Biancone, L., Araki, M., Araki, K., Vassalli, P. & Stamenkovic, I. Redirection of tumor metastasis by expression of E-selectin in vivo. J. Exp. Med. 183, 581–587 (1996).
61. Herschkowitz, J. I. et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. 8, R76 (2007).

Acknowledgements
We thank L. M. Wakefield for providing the SORE6–mCherry stemness reporter, R. Nusse for providing the 7×TCF–GFP Wnt reporter, G. Laevsky for assistance with microscopy and C. DeCoste for assistance with flow cytometry. This work was supported by fellowships from the NIH (grant no. F31CA192461) and NJCCR to M.E., the National Institutes of Health NHLBI (grant no. P01 HL107146), the Program of Excellence in Glycosciences and the Team Joibe Fund to R.S., and grants from the Susan G. komen Foundation (grant no. SAC160867), Glycomimetics Inc., Brewster Foundation, Department of Defense (grant no. BC123187) and the National Institutes of Health (grant no. R01CA141062) to Y.K. This research was also supported by the Preclinical Imaging, Genomic Editing and Flow Cytometry Shared Resources of the Rutgers Cancer Institute of New Jersey (grant no. P30CA072720).

Author contributions
M.E. and Y.K. conceived the project and co-wrote the manuscript. M.E. designed and conducted all of the flow cytometry, xenograft, genetic, qRT–PCR, confocal and bioinformatic experiments, and analysed the data. M.E. derived the SUM159-M1a and DU145-ob1 cell lines. T.M.G., M.E. and I.M.C. performed the mass spectrometry experiments, provided advice and assisted in writing the manuscript. Y.W., C.S., H.Z. and C.C. assisted with the mouse experiments and stable cell generation. Y.W. and M.E. performed and analysed the microarray experiments. S.-C.L. and S.-H.L. stained the prostate-cancer bone biopsies. J.L.M. provided GMI-1271 and research input.

Competing interests
J.L.M. is the Vice President and Chief Scientific Officer of Glycomimetics, Inc., which owns the patent to GMI-1271. Y.K. received research support from Glycomimetics Inc. for experiments using GMI-1271. No other authors declare any conflicts of interest.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41556-019-0309-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to Y.K.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019
Clone sequences flanked by EcoR1 and Hpa1 restriction sites were inserted into the pMSVC retroviral plasmid. The clones were sequenced and compared with NCBI expressed sequence tags for accuracy. Catalytic mutants of Fut3 were generated by DhaII digestion and ligation into pBluescript KSII plasmid. Mutations were confirmed by DNA sequence analysis. The catalytic domain was then excised by EcoR1 and SspI digestion and ligated into the pMSVC retrovirus plasmid. The resulting plasmids were transfected into 293T cells and the supernatants were collected 3 days later. The retroviruses were harvested and filtered through 0.2 μm pores and spun at 4000 x g for 30 minutes. The resulting supernatants were used to infect the cells.

Flow cytometry. For E-selectin binding experiments, a RFP-labelled internal control line and GFP-labelled cells were co-cultured at equal ratios to 48 h; cells were then harvested at 80% confluence by non-enzymatic dissociation buffer (Life Sciences) at 37 °C, washed once with PBS and incubated with anti-human IgG–ALE647-conjugated antibody (BioLegend) at 0.1 μg/ml for 30 min. The cells were then washed and analyzed by flow cytometry using the BD LSRII flow cytometer. For Western blotting, cells were lysed in RIPA buffer and the lysates were separated by SDS-PAGE and transferred to PVDF membranes. The membranes were blocked in 5% milk in TBS–T, probed with antibodies listed in Supplementary Table 8 and imaged on the Licor Odyssey CLX system using Licor-supplied IRDye680 and IRDye800 secondary antibodies. Dkk1–FLAG was obtained (Addgene, cat. no. 16690) and subcloned into the pCMV vector for lentiviral production. Knockdowns were performed using lentiviral shRNA vectors purchased from Sigma (Supplementary Table 7) and knockouts performed using the lentiviral CRISPR–Cas9 vector system pLentiCrisp-v2 (Supplementary Table 7). Viruses were produced in HEK293FT cells using the pLVX packaging system (Clontech) and used at MOIs of 1.5 to achieve efficient transduction.

Cell surface protein isolation. Cell surface proteins were biotinylated using the Pierce cell surface protein isolation kit (Thermo Scientific). Briefly, WT and variant M1a cell lines were harvested from two 175 cm2 flasks by trypsinization, rinsed with PBS, and suspended in flow buffer. The cells were cultured with 10% FBS, 10 μM Dkk1–FLAG was added, and incubated for 12 h, followed by culture with anti-Slug, which was detected with the Odyssey CLX system using Licor-supplied IRDye680 and IRDye800 secondary antibodies. Anti-Flag was detected with the Chemifluorescent substrate kit (LicoR). E-selectin immunoprecipitation western blots were performed in the presence of 1% sodium dodecyl sulfate (SDS), for E-selectin–IgG and, along with anti-Slug, were detected with traditional horseradish peroxidase detection. The cell proliferation rates were quantified using the EZQuant reagent (Altem Bios).

Cell surface protein isolation. Cell surface proteins were biotinylated using the Pierce cell surface protein isolation kit (Thermo Scientific). Briefly, WT and variant M1a cell lines were harvested from two 175 cm2 flasks by trypsinization, rinsed three times with PBS and suspended in 0.25 mg/ml solution of Sulfo-NHS-SS-Biotin in 100,000 cells/ml. The cells were then washed two times with PBS and resuspended in 100 μl quenching solution, which was added to each tube followed by centrifugation at 1,000 rpm for 5 min. The cell pellet was then washed twice in 10 ml Tris-buffered saline. The cell pellets were then lysed with buffer containing 1% NP40. The lysates were equilibrated with Neutra Avidin agarose beads for 2 h at 4 °C. Cell membrane
proteins were eluted by boiling the beads with 100 μl western blot sample buffer with 50 mM DTT for 5 min at 95°C. 

Immunoprecipitation with E-selectin–IgG. M1a cell variants (5 × 10⁶) were harvested from tissue culture flasks and washed three times with PBS. The cell lysates were prepared in Selectin wash/lysis buffer (2% NP40, 150 mM NaCl, 50 mM Tris–HCl (pH 7.4), 2 mM CaCl₂, 20 μg/ml PMSF, and 1μM protease inhibitor cocktail (Roche)). The cell lysates were pre-cleared by incubating overnight with Protein G agarose beads (Invitrogen) pre-blocked with BSA. The pre-clear lysates were then mixed with 10 μg/ml E-selectin–Ig for 4 h at 4°C. E-selectin–Ig binding proteins were pulled down via incubation with Protein G agarose beads pre-blocked with BSA and eluted through boiling in 1.5X reducing sample buffer with SDS.

Immunofluorescence. Ex vivo immunofluorescence on bone, lung and liver samples was conducted with or without retro-orbital injection of 10 μg anti-C1D31 (Biolegend, cat. no. 102416) in 100 μl PBS to label the vasculature, followed by whole-body perfusion with 1% buffered formalin. The tissues were removed and fixed for 12 h with 10% buffered formalin at 4°C. These were equilibrated in 15 and 30% sucrose and incubated in 2% PFA (R & D Systems, B8491; 1:50) at 4°C followed by Alexa Fluor-conjugated secondary antibody (1:300). The sections were then sequentially incubated with anti-Glg1 (Sigma-Aldrich, HPA010815; 1:100) for 12 h in 20% goat serum and 0.1% Triton X-100, and the nuclei were counterstained with DAPI. Images were taken on an Olympus Confocal

Immunoprecipitation with E-selectin–Ig for 4 h at 4°C. E-selectin–Ig binding proteins were pulled down via incubation with Protein G agarose beads pre-blocked with BSA and eluted through boiling in 1.5X reducing sample buffer with SDS.

Microarray analysis. BM2 and M1a cells were seeded over 6 cm tissue-culture-treated plates coated with E-selectin–Ig or control IgG (10 μg/ml−1) for 12 h at 37°C. RNA was isolated from cells using the RNAeasy minikit (Qiagen) according to the manufacturer’s instructions. Samples were amplified using thephereo QuickAmp PCR system (Agilent, G4858A-039494). Briefly, the RNA samples and a universal human reference RNA (Agilent) were labelled with CTP-cy5 and CTP-cy3 using the Agilent Quick Amp labelling kit. Labelled samples were mixed equally and hybridized to the array. The array was then scanned with the G2505C scanner (Agilent) and analysed with the Agilent GeneSpring GX software (Agilent). Briefly, array controls, flagged values and expression values falling below the median value were removed. Multiple values for any given gene were collapsed into the single highest expression value. Data was extracted as a log2-transformed ratio of Cy5/Cy3.

GSEA. Data that was log-transformed was subtracted (log2[E-selectin]−log2[IgG]) and rank ordered for each cell line. The data was analysed using GSEA v2.0. Intergroup signatures from the MySigDB v5.1 database included the Hallmark EMT dataset (M5930), Sarrio–EMT (GSE84330), Claudin-low (GSE18229), Luminal (GSE22446), Bruno-hematopoiesis (M1492) and Blick-EMT–5 as part of a broader, manually compiled set of EMT- and Stem cell-related signatures. Only data sets with a nominal P < 0.05 were considered in the analysis.
Data availability
The microarray data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under the accession code GSE96754 and the mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012942; the mass spectrometry data is further available in Supplementary Table 2. The human clinical breast cancer data are not available in GEO; these were derived from the TCGA Research Network and the dataset derived from this resource that supports the findings of this study is available in https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Breast%20Cancer%20(BRCA). The human breast cancer data were also derived from the NKI-295 dataset\(^6\), which is available at https://xenabrowser.net/datapages/?cohort=Breast%20Cancer%20. Finally, the human breast cancer data from the EMC-MSK dataset is available in the publication by Bos and colleagues\(^6\). The source data relevant for the clinical data analyses performed in this study are available in Supplementary Table 9. Unprocessed western blot images are provided as Supplementary Fig. 9. The source data supporting the findings of this study are provided in Supplementary Table 9. All protocols, cell lines and reagents are available from the corresponding author on request.

References
62. Forozan, F. et al. Molecular cytogenetic analysis of 11 new breast cancer cell lines. Br. J. Cancer 81, 1328–1334 (1999).
63. Ell, B. et al. Tumor-induced osteoclast miRNA changes as regulators and biomarkers of osteolytic bone metastasis. Cancer Cell 24, 542–556 (2013).
64. Koda, Y., Kimura, H. & Mekada, E. Analysis of Lewis fucosyltransferase genes from the human gastric mucosa of Lewis-positive and -negative individuals. Blood 82, 2915–2919 (1993).
65. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. Nat. Methods 11, 783–784 (2014).
66. Diner, B. A. et al. The functional interactome of PYHIN immune regulators reveals IFIX is a sensor of viral DNA. Mol. Syst. Biol. 11, 787 (2015).
67. Giguere, S. S. et al. The proteomic profile of deleted in breast cancer 1 (DBC1) interactions points to a multifaceted regulation of gene expression. Mol. Cell. Proteomics 15, 791–809 (2016).
68. Blick, T. et al. Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44\(^+\)/CD24\(^-\) stem cell phenotype in human breast cancer. J. Mammary Gland Biol. Neoplasia 15, 235–252 (2010).
69. Bos, P. D. et al. Genes that mediate breast cancer metastasis to the brain. Nature 459, 1005–1009 (2009).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  - Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Data was collected using standard software programs including: Nikon Elements for microscopy, FlowJo v10 and FACSDiva for flow cytometry, Licor Odyssey Clx for western blot, ImageJ for bone metastasis x-rays, LivingImage for bioluminescence, Skyline for Mass spectrometry, Genespring v13 for microarray, and GSEA-BroadInstitute for GSEA analysis. These have mentioned in the text where its reporting is customarily used. Stata and Excel were used for basic data display and statistical tests. No code was used to collect any data.

Data analysis

Excel functions ttest and stdev used for Student’s T-test and standard deviation calculations. Stata ranksum used for Mann-Whitney U test and stcox used for Cox’s Proportional Hazards model. Vioplot add-in used for stata graphing of violin plots. GSEA pre-ranked GSEA used for GSEA analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Microarray data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession codes GSE96754, and the mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012942; mass spectrometry data is further available in Supplementary Table 2.

The human clinical breast cancer data are not available in GEO; these were derived from the TCGA Research Network and the data-set derived from this resource that supports the findings of this study is available in https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Breast%20Cancer%20(BRCA)

The human breast cancer data were also derived from the NKI-295 dataset which is available at https://xenabrowser.net/datapages/?cohort=Breast%20Cancer%20NKI-295.
Finally, human breast cancer data from the EMC-MSK dataset is available in Bos et al., 2009. Source data relevant for the clinical data analyses performed in this study are available in Supplementary Table 9.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | For most experiments, biological replicates and/or technical repeats were included and most experiments were repeated at least 3 times. Such sample sizes are typical of the in vitro experiments (e.g. qRT-PCR) used in the study. For in vivo experiments, sample size of n = 8-10 are typical of experiments and are used in the study, but this number varies according to the number of wild-type and knockout mice per litter in the E-selectin knockout experiment or when the number of mice in a single experiment was >40 animals. No sample size calculation was performed to predetermine sample size. All sample sizes are listed in the corresponding figure legend. |
| Data exclusions | Animals were excluded only if they died or had to be sacrificed because of moribund conditions specified in the IACUC protocol. Such events occurred in less than 5% of animals during the experimental time frame. This exclusion criteria was pre-determined by the protocol. |
| Replication | Immunofluorescence, western blot, mass spectrometry controls, flow cytometry and E-selectin plate-binding assays were conducted ≥3 times to ensure reproducibility unless otherwise noted. All in vivo bone metastasis experiments were performed at least 2 times to ensure reproducibility and significance with the exception of the BM2-CD44 KO injection, catalytic Fut3 mutants and the GMI-1271 treatment, which was performed once. All lung metastasis and primary tumor growth experiments were performed once with appropriate group size and statistics. All attempts at replication were successful. |
| Randomization | Standard mice (Nu/Nu or NSG) for xenograft experiments were randomized, with each group receiving an equal number of mice from each litter. E-selectin knockout compared to wild-type mice were not randomized, rather, every available mouse within 20 days of the same date of birth was used. |
| Blinding | No blinding was performed with the exception of x-ray analysis and quantification. In this case, the investigator was blinded to the experimental group of each x-ray. For all other cases, blinding was not relevant to the experiment as the results were directly collected by instrumentation. |

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Study description | Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study). |
| Research sample | State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source. |
| Sampling strategy | Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed. |
| Data collection | Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection. |
| Timing | Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort. |
| Data exclusions | If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established. |
| Non-participation | State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation. |
Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

- **Study description**: Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.

- **Research sample**: Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxonomy, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.

- **Sampling strategy**: Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.

- **Data collection**: Describe the data collection procedure, including who recorded the data and how.

- **Timing and spatial scale**: Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken.

- **Data exclusions**: If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.

- **Reproducibility**: Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.

- **Randomization**: Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.

- **Blinding**: Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

- **Did the study involve field work?**: Yes ☐  No ☐

**Field work, collection and transport**

- **Field conditions**: Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).

- **Location**: State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).

- **Access and import/export**: Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).

- **Disturbance**: Describe any disturbance caused by the study and how it was minimized.

### Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

| Materials & experimental systems | Methods |
|----------------------------------|---------|
| n/a | Involved in the study | n/a | Involved in the study |
| ☑ | Antibodies | ☑ | ChIP-seq |
| ☑ | Eukaryotic cell lines | ☑ | Flow cytometry |
| ☑ | Palaeontology | ☑ | MRI-based neuroimaging |
| ☑ | Animals and other organisms | | |
| ☑ | Human research participants | | |
| ☑ | Clinical data | | |
Antibodies

Antibodies used

- Glg1 (Sigma HPA010815, 1:1000-IB or 1:100 IF), CD44 (Cell Signaling Technologies clone 156-3C11, 1:1000), Beta-actin (Santa Cruz sc7778, 1:1000), N-cadherin (BD Pharmigen 610920, clone 32, 1:1000-IB, 1:100 IF), FLAG (Sigma F7425, 1:1000), Fut3 (Abcam ab110082, 1:1000), E-selectin (BD Pharmigen clone 109.E6, 1:100), EpCam (CST VU1D9, 1:100), E-cadherin (BD Pharmigen 610181, clone 10E6, 1:100), Ki67 (Dako M7240, clone MIB-1, 1:40), ZO-1 (CST S406, 1:100), EpCam (CST VU109, 1:100), K14 (Biologend 905301, 1:100), Slug (Santa Cruz clone A-7, 1:50), E-selectin (R&D systems BBA18, 1:50), Glg1 (Personal gift, Charles Dimitroff, 1:1000). This data is available in Supplementary Table 8.

All secondary antibodies used were Licor iRdye 680/800 anti-rabbit or anti-mouse for western blot (925-322#, 1:10000), Dako polyclonal anti-rabbit or anti-mouse-HRP for western blot (PO448, 1:2000) or Invitrogen ALS55 or AL647 anti-rabbit or anti-mouse (e.g. A-21244, 1:1000) for immunofluorescence.

Validation

- All commercial antibodies were validated by the manufacturer for the species and application used in this study. Glg1, CD44, Fut3, and FLAG antibodies were validated for western blot via specific knockout or overexpression followed by western blotting- data for each is displayed in the manuscript. All bands appeared at the correct apparent molecular weight. Beta-actin, Slug, CD31, EpCam, E-cadherin, K-14, Ki-67 antibodies were previously validated in our lab for the intended species and application used here.
- E-selectin and Glg1 antibodies were validated for IF by overexpression in HEK-293T cells followed by indirect immunofluorescence.
- In addition, all antibodies used here have been validated by reviews on Biocompare. anti-Glg1 from Charles Dimitroff has been previously validated by Robert Sackstein’s laboratory and was furthermore validated in this study by specific knockout and overexpression of Glg1 followed by western blotting in Supplementary Fig. 5.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

- Key characteristics of each cell line, including source, are included in Supplementary table 1. The MDA-MB-231 cells and its derivatives (SCP6, SCP28, SCP2, SCP25, LM2, BM2) were from Joan Massague’s lab. SUM159 was from Stephen Ethier’s lab at the Karmanos Institute and the derivatives were made by this study’s lead author. DU145 was obtained from ATCC and the ob1 bone-metastatic derivatives were generated by this study’s lead author. L Cells were obtained from ATCC. PCR1 cell line was obtained from Steven Barthel.

Authentication

- All cell lines were authenticated by STR profiling with a similarity score >90%.

Mycoplasma contamination

- All cultures were confirmed negative and no instances of mycoplasma contamination occurred throughout the study period.

Commonly misidentified lines

(See ICLAC register)

- No cell lines used in this study were found in the database of commonly misidentified cell lines that is maintained by ICLAC and NCBI Biosample.

Palaeontology

Specimen provenance

- Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

Specimen deposition

- Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

- If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

Animals and other organisms

Policy information about studies involving animals, ARRIVE guidelines recommended for reporting animal research

Laboratory animals

- E-selectin knockout mice were ordered from the Jackson Laboratory (B6.129S4-Selemtm1/DmiL) and were backcrossed for 5 generations into the NOD/SCID strain. All xenograft experiments were conducted on 8 week old female mice (athymic Nu/Nu, NOD/SCID, NOD/SCID Gamma) except in prostate cancer xenografts, which were performed on male mice of the same age. All mice were originally ordered from the Jackson Laboratory and breeding was conducted in an SPF barrier facility.

Wild animals

- The study did not use wild animals.

Field-collected samples

- The study did not use field-collected samples.

Ethics oversight

- All procedures involving mice and experimental protocols were approved by the University Institutional Animal Care and Use Committee (IACUC).
Human research participants

Policy information about studies involving human research participants

| Population characteristics | Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above." |
| Recruitment | Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results. |
| Ethics oversight | Identify the organization(s) that approved the study protocol. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

| Clinical trial registration | Provide the trial registration number from ClinicalTrials.gov or an equivalent agency. |
| Study protocol | Note where the full trial protocol can be accessed OR if not available, explain why. |
| Data collection | Describe the settings and locales of data collection, noting the time periods of recruitment and data collection. |
| Outcomes | Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures. |

ChIP-seq

Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as GEO.
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

| Replicates | Describe the experimental replicates, specifying number, type and replicate agreement. |
| Sequencing depth | Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end. |
| Antibodies | Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number. |
| Peak calling parameters | Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used. |
| Data quality | Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment. |
| Software | Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details. |
Flow Cytometry

Plots

Confirm that:

☑️ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
☑️ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).
☑️ All plots are contour plots with outliers or pseudocolor plots.
☑️ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

For E-selectin binding experiments, RFP- and GFP-labeled cells were co-cultured at equal ratios for 48 h, then cells were harvested at 80% confluence by non-enzymatic dissociation buffer (Life Sciences) at 37°C, washed once with PBS, and suspended in flow buffer (10% FBS in PBS supplemented with 1 mM Ca2+/Mg2+) at 1 million cells/mL. Recombinant mouse E-selectin/IgGFc or Isotype IgG (RD systems) was added at 10 ug/mL for 1.5 h with vortexing every 15 min. Cells were washed with PBS and incubated with anti-human IgG-AL647-conjugated antibody (Biolegend) at 0.1ug/mL for 45 min. Cells were washed with PBS, resuspended in flow buffer with DAPI (1 ug/mL), and analyzed with the BD LSRII flow cytometer. Internal control cell lines (RFP-labeled) corresponded to the parental population of each cell line – either MDA-MB-231-RFP or SUM159-RFP. E-selectin binding ratios were quantified using the formula (SELEGFP/IgGGFP)/(SELERFP/IgGRFP). Gating for DAPI, RFP, and GFP were performed using negative controls of the corresponding cell line using Flowjo version X. For pharmacological, cells were grown in Tunicamycin, Deoxymannojirimycin, or PDMP for 48h. Following differential treatment, internal control RFP-labeled cells were added to the cell mixture. For FACS sorting, MDA-MB-231 cells were labeled using this protocol and were sorted into the top and bottom 10% of binding intensities. E-selectin-sorted MDA-MB-231 cells were passaged 5 times prior to assessing E-selectin binding levels to generate a large enough pool both for flow cytometry and sub-culturing. Negative control treatments (EDTA or GMI-1271) were tested in every cell line to ensure that E-selectin binding was specific. For Wnt signaling activation analysis, BM2 cells stably transduced with the 7x-TCF-GFP-SV40-mCherry (TGC) reporter were trypsinized, washed once with PBS, and suspended in PBS + 1 ug/mL DAPI. Cells were analyzed with the LSRII instrument, and gating analysis was performed using unlabeled BM2 cells using the Flowjo version X software. All flow cytometry experiments were repeated a minimum of three times.

Instrument

BD LSRII

Software

FACSDiva and FlowJo Version X were used to set gates and collect data.

Cell population abundance

No cell sorting was performed beyond isolating the top and bottom deciles for the E-selectin-sorted cell lines.

Gating strategy

FSC and SSC gates were set to include at least 90% of events and to exclude very small events. SSC-A and -H or FSC-A and -H were used to discriminate for only single cells. DAPI or PI were used to discriminate live cells from dead using appropriate negative and positive controls. GFP and RFP gating was performed using negative controls from the respective cell lines tested, or from using appropriate internal controls (i.e. the Sore6 reporter system which used a control plasmid to control for endogenous mCherry transcription) such that the negative control sample showed no more than 1% positivity. All binding analyses were performed using an internal control to control for differences in cell culture, harvesting or probing conditions. This internal control was used to normalize all experiments as described.

☑️ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).
### Acquisition

| Imaging type(s) | Specify: functional, structural, diffusion, perfusion. |
|----------------|------------------------------------------------------|
| Field strength | Specify in Tesla                                      |
| Sequence & imaging parameters | Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle. |
| Area of acquisition | State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined. |
| Diffusion MRI | □ Used □ Not used |

### Preprocessing

| Preprocessing software | Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.). |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Normalization | If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization. |
| Normalization template | Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized. |
| Noise and artifact removal | Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration). |
| Volume censoring | Define your software and/or method and criteria for volume censoring, and state the extent of such censoring. |

### Statistical modeling & inference

| Model type and settings | Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation). |
|-------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Effect(s) tested | Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used. |
| Specify type of analysis: □ Whole brain □ ROI-based □ Both | Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods. |
| Statistic type for inference (See Eklund et al. 2016) | Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods. |
| Correction | Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo). |

### Models & analysis

| n/a | Involved in the study |
|-----|-----------------------|
| □   | □ Functional and/or effective connectivity |
| □   | □ Graph analysis |
| □   | □ Multivariate modeling or predictive analysis |

| Functional and/or effective connectivity | Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information). |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Graph analysis | Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.). |
| Multivariate modeling and predictive analysis | Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics. |