SUPPLEMENTAL METHODS

Ethics Statement

Human archival tissue was obtained from the Ontario Institute for Cancer Research (OICR), Ontario Tumor Bank. All experiments involving animals were approved by the Animal Use Subcommittee at the University of Western Ontario (Protocol No. 2008-101).

Immunohistochemistry (IHC) of human tissue

Formalin-fixed, paraffin-embedded archival tissue with clinical and pathological information was obtained from breast cancer patients (OICR). In total, ninety tumor sections, predominately reflecting invasive disease, were procured. Following deparaffinization in xylene, ethanol degradation, antigen retrieval with citrate buffer, and peroxidase and serum-free protein blocks, Nodal or CD31 specific antibodies were applied (**Supplementary Table 1**). To verify the specificity of Nodal staining, a monoclonal versus a polyclonal antibody were compared (**Supplementary Figure 1)**. In confirmation with previous studies, these antibodies stained the same cells, and revealed comparable staining patterns (13). We proceeded with the monoclonal antibody, as it generated less background staining, has been used successfully in several publications, and has been shown to detect human Nodal in breast cancer, melanoma and endometrial cancers (5;6;14). Slides were rinsed in TBS-T, and treated with Envison+ HRP anti-mouse IgG (Dako). Color was produced with DAB substrate and counterstained with Mayer's haematoxylin. Samples were dehydrated in reagent grade alcohol and cover slipped with permanent mounting medium. Negative controls were conducted with mouse IgG, isotype-matched and used at the same concentration as Nodal and CD31 antibodies.

Evaluation of Nodal and vascular density in breast cancer sections

Nodal IHC was evaluated under light microscopy. Cytoplasmic staining was seen in all cases. Cases were blindly and independently scored by two pathologists (S.J.D. and J.M.) to derive an Allred Score (15). Total Allred score is the sum of proportion score (percentage of cells stained: 0%=0, <1%= 1, 1-10%= 2, 10-33%= 3, 34-66%= 4 and >67%= 5) and intensity score (No staining= 0, Weak= 1, Moderate= 2, Strong= 3). Values were dichotomized using a total Allred score of 7 and 8 as high expression, and ≤6 as low expression. Vascular density was blindly and independently scored by two pathologists (S.J.D. and J.M.) based on CD31 IHC staining, following the recommendations from the International Consensus on Evaluation of Angiogenesis in Solid Human Tumors (16). In summary, three "hotspots" in the stromal component were selected, the number of vessels in each "hotspot" was counted at 200x magnification, and a final score was expressed as the mean vessel density. Data was dichotomized using the average number of vessels among all samples as the cut-off between high and low vascular density values. All scoring values obtained by S.J.D. and J.M. had a Pearson Correlation Coefficient (rp) of 0.73 or higher. 83 cases out of 90 cases had both Nodal IHC and CD31 vascular density scores, and were used for statistical analyses. For clinical characteristics of these 83 patients, see **Supplementary Table 2**. The Allred scores for Nodal were used to assess correlations between Nodal, ER status, PR status, HER-2 status and tumor grade. For these correlation analyses, positive ER/PR/HER-2 staining was assigned a value of 1 and negative staining was assigned a value of 0; and the disease grade was assigned a value of 1, 2 or 3. Assessment of correlations was performed using Pearson Correlation and a 2-sided t-test.

Cell culture and transfection

Endothelial cells: Two endothelial cell lines were used, human umbilical vein endothelial cells (HUVECs; ATCC), and primary human adult microvascular endothelial cells (HMVECs; Invitrogen). We used both adult and umbilical sources to account for possible functional differences in response to Nodal, since Nodal is an embryonic protein. Cells were maintained with attachment factor, and Medium 131 + Microvascular Growth Supplement as per manufacturer suggestions (Invitrogen).

Breast cancer cells: We used two highly aggressive breast cancer cell lines that express high endogenous Nodal for loss-of-function studies (MDA-MB-231 and MDA-MB-468), and one poorly aggressive breast cancer cell line that expresses low endogenous Nodal for gain-of-function studies (T47D). All breast cancer cell lines were obtained from and validated by ATCC, and were maintained as per instructions. Multiple constructs were used to knock down or induce Nodal signaling in each of the cell lines used:

MDA-MB-231 cells:

- GIPZ lentiviral shRNAmir against Nodal 3rd exon (Clone: V2LHS_155453, Open Biosystems).
- TRIPZ lentiviral shRNAmir against Nodal 3rd exon (Clone: V2THS_155453) with Tet-On®/Tet-Off® inducibility (Open Biosystems)

MDA-MB-468 cells:

- HuSH-29mer shRNAs against Nodal 3rd exon (Id: GI311711) (Origene)

T47D cells:

- A Nodal expression vector was made with pcDNA™3.3-TOPO® cloning kit, and an empty pcDNA3.3 vector was used as a control. Vectors were sequenced and validated.

- Recombinant human Nodal (50-100 ng/mL) was used to treat cells and activate P-SMAD2 signalling (R&D systems).

For all knockdown systems, shRNAs targeting at least 4 regions in the Nodal gene were tested for their ability to knockdown Nodal protein expression. For each vector type, the shRNA achieving the best knockdown was chosen. Stable control and anti-Nodal shRNAs were transfected using Arrest-In (Open Biosystems) or Lipofectamine (Invitrogen) as per manufacturer instructions, and cells were selected using Puromycin (200-450 ng/mL).

RNA extraction and Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)

Total RNA was isolated from cells and tumour tissue using the Perfect Pure RNA cultured cell kit (5 Prime) and genomic DNA was degraded using DNAse. Reverse transcription was performed using a High Capacity cDNA Reverse Transcription kit (Applied Biosystems). Real-time PCR was performed using 2 μg cDNA with TaqMan® gene expression human primer/probe sets (**Supplementary Table 3)**. Gene expression was normalized to the endogenous control genes HPRT1 or RPLPO.

Western blotting

Protein lysates were prepared and quantified as previously described (6). Equal amounts of protein were separated by SDS-polyacrylamide gel electrophoresis under reducing conditions, and resolved proteins were transferred onto Immobilon-P membranes (Millipore Corp.). Membranes were incubated with primary antibody (**Supplementary Table 1**) and the appropriate horseradish peroxidase-labelled secondary antibody. Secondary antibodies were detected by enhanced chemiluminescence (Super Signal; Pierce). For Nodal, in accordance with

previous studies (5;6;17;18), up to 3 species were detected: A pro-Nodal band at ~39 kDa, multiple processed (glycosylated etc.) bands at ~ 50 kDa, and a cleaved Nodal band at ~15kDa. These bands have been shown to be proportionally expressed (18); however, they vary with respect to ease of detection. The 50 kDa bands are highly variable due to differences in post-translational modifications and protein lysate handling. The 15 kDa band is poorly abundant and is most readily observed in conditioned media. For consistency, and accuracy we used the 39 kDa band to assess Nodal expression in lysates and the 15 kDa band to assess Nodal levels in conditioned media.

In vitro functional assays

Functional assays were performed on endothelial cell lines (HUVECs and HMVECs) in response to conditioned media from either (1) MDA-MB-231 cells transfected with a control shRNA or a Nodal-targeted shRNA, or (2) T47D cells transfected with an empty control vector or a Nodal overexpression vector.

Tube formation: Endothelial cell Medium 131 supplemented with 0.5% BSA was conditioned by cancer cells for 24 hours. Endothelial cells were treated with conditioned media and tube formation was quantified. In some cases, rhNodal (100 ng/mL; R&D), VEGF (20 ng/mL; Invitrogen), and/or PDGF (20 ng/mL; Invitrogen) were added to the conditioned medium *after it was collected*, during the course of the assay. 20 ng/mL of VEGF and PDGF was used because it was both the minimum concentration consistently reported to stimulate endothelial cells *in vitro* (19;20), and because it is much lower than concentrations used for most *in vivo* positive controls (for example, refer to DIVAA manufacturer protocol). Tube formation was

quantified by counting the number of circular structures that formed from complete branching per field of view. A sum of 10 fields of view was used for each treatment replicate.

Migration: Endothelial cells (50,000 cells) were seeded on a Transwell insert, and treated with conditioned media from cancer cells. Migration (24 hrs) was quantified by staining nuclei with DAPI and counting cells. In some cases, rhNodal, VEGF, and/or PDGF were added to the conditioned medium *after it was collected*, during the course of the assay. An average of 10 fields of view (20x magnification) was used for each treatment replicate.

Viability: Viability of endothelial cells in response to conditioned media was assessed using the LIVE/DEAD® Viability/Cytotoxycity Kit as per manufacturer's instructions (Invitrogen).

Proliferation: Cellular proliferation was quantified with a haemocytometer and Trypan Blue, and by automatic counting with a *Countess* (Invitrogen) cell counter and Trypan Blue.

Measurement of angiogenic proteins

All protein measurements for VEGF and PDGF were performed using Aushon SearchLight Chemiluminescent Angiogenesis Arrays (#84694; Aushon Biosystems, Billerica, MA) as per manufacturer instructions, or by Western blot analysis.

Directed In Vivo Angiogenesis Assay (DIVAA)

A DIVAA kit was used for the *in vivo* angiogenesis assays (Trevigen, Gaithersburg, MD). Angioreactor tubes were filled with matrix (negative control), matrix + cancer cells, or matrix + cancer cells + growth factors (for rescue analyses). As in our *in vitro* assays, rhNodal was added at a concentration of 100 ng/mL, and VEGF was added at a concentration of 20 ng/mL.

Angioreactors were inserted subcutaneously into nude mice. After 10 days, mice were sacrificed and angioreactors were removed. Blood vessels were imaged under a microscope, and contents of angioreactors were transferred into a centrifuge tube. Endothelial cells were labelled with FITC-Lectin, and fluorescence was quantified.

Chorioallantoic Membrane (CAM) tumor angiogenesis assay

One collagen-enmeshed grid containing 2 X 10^6 cells was placed on each CAM, and newly formed blood vessels that appeared in the upper grids of the collagen onplants were scored and quantified as previously described (21). Tumor cells were visualised with GFP expressed in the bi-cistronic GIPZ shRNA constructs. Recombinant VEGF was added at a concentration of 20 ng/mL for rescue analysis.

In vivo inducible shRNA tumor assay

MDA-MB-231 cells were transfected with a Doxycyclin-inducible Control or Nodal-targeted shRNA. 500,000 cells in RPMI + Matrigel (1:1; 100 µL) were injected into the right flank of 6-8 week old female Nude mice (Crl:NU-*Foxn1*^{nu}*;* Charles River). Doxycyclin was administered 2 weeks following onset of palpable tumor growth via diet (0.625 g doxycyline hyclate per 1 kg chow). This diet is optimized to deliver Doxycyclin at 2-3 mg/day, given that each mouse eats approximately 4-5 g/day (Harlan Laboratories). Excised tumors were cut in half: one half formaldehyde-fixation for IHC, other half frozen and cryosectioned for IF and/or RNA extraction for PCR.

Immunofluorescence: Immunofluorescence was conducted using anti-CD31 and anti-Nodal antibodies (**Supplementary Table 1**) with the appropriate fluorochrome-conjugated

secondary antibodies. DAPI was used to stain nuclei and IgG isotype controls were performed. For each tumor section, CD31 was quantified by taking 5 random images of vascular hotspots around the tumor periphery (tumor cores were too necrotic to include), and then counting the number of vessels in each field of view across a 25-box grid overlay using ImageJ software. For each tumor, 3 serial sections were quantified by this method, and averaged to yield one value for CD31 expression per mouse replicate.

Analysis of tumor necrosis: Three tumor sections spaced evenly throughout each tissue block were stained with H&E. Each section was imaged such that the entire section was visible in one picture. Photoshop software was used to outline and quantify total tumor area, and area of necrosis. Necrosis was calculated as a percentage of the total tumor area.

Analysis of tumor hypoxia: Animals were injected with hypoxyprobe-1 (60 mg/kg body weight) prior to sacrificing. Excised tumors were frozen and IHC was performed as described above, with a few modifications. Acetone was used for fixation. Endogenous immunoglobulins in mouse tissue were blocked with the mouse-on-mouse (M.O.M.™) basic kit (Vector Laboratories). Sections were incubated with anti-hypoxyprobe-1 antibody (HPI, Inc) (**Supplementary Table 1**), and a biotinylated anti-mouse IgG (H+L) secondary antibody (M.O.M.™ kit, Vector Laboratories). Vectastain ABC kit (Vector Laboratories) was used to detect enzyme activity, DAB substrate (Dako) was used for color development, and haematoxylin was used to counterstain.

Statistical analyses Statistical analyses of multiple comparisons for parametric data were performed using a one-way ANOVA followed by a Tukey–Kramer Comparisons Post-Hoc test**.** An ANOVA on Ranks followed by the Mann-Whitney rank-sum test was used for non-

parametric data. When only two items were compared, a student's t-test was used. Parametric data were expressed as mean ± S.E.M. for replicate values, and non-parametric data were expressed as median ± interquartile range. All statistical tests were two-sided and data comparisons for all experiments were considered statistically significant at $p<0.05$. Statistics were performed using SigmaStat (Dundas Software) in consultation with the biostatistical support unit at the University of Western Ontario.