Supplementary Materials and Methods

Patient Information and Tissue Collection

Histologically, 3/30 primary tumors were classified as Grade 1, whereas the remaining 27 cases were Grade 3. Pathological staging confirmed the presence of lymph node involvement in 15/30 patients. Metastatic lesions (histologically diagnosed) were also obtained from six breast cancer patients. No patient received neoadjuvant chemotherapy or radiation therapy prior to surgery and tissue collection.

Immunohistochemical Detection of Brachyury and Scoring Method

Sections of paraffin-embedded, formalin-fixed tissues were evaluated for brachyury (T, brachyury homolog) expression as previously described (1). Briefly, tissue sections were deparaffinized in xylene, rehydrated in a series of graded ethanol, and treated with 0.3% H2O2 in methanol to block endogenous peroxidase activity. Microwave-citrate buffer antigen retrieval method was performed to unmask the antigen. The sections were blocked in 10% horse serum (Invitrogen, Carlsbad, CA) for 1 hour at room temperature and then incubated overnight at 4°C with a mouse anti-brachyury monoclonal antibody (Ab) (ab57480, Abcam, Cambridge, MA) at a 1:100 dilution. In addition, a positive control Ab (mouse anti-Cytokeratin Ab) and an isotype matched Ab (MOPC 21, Sigma-Aldrich, St. Louis, MO) were used to verify accurate staining method. Immunostaining was carried out using the Vectastaining ABC kit (Vector Laboratories, Burlingame, CA) following the manufacturer's instructions; color was developed with 3,3′-Diaminobenzidine (DAB) peroxidase substrate. Sections were counterstained with haematoxylin, dehydrated in ethanol, cleared in xylene, and mounted under a coverslip using Permount (Fisher Scientific, Fair Lawn, NJ).

Two pathologists independently evaluated the tumor and normal tissue samples in a blinded, randomized way. For each slide, three to five random fields were evaluated; for each field, the percentage of DAB-positive-tumor cells was calculated as: [(number of DAB-positive tumor cells/total number of tumor cells) x 100], and the relative staining intensity was scored as weak (+) for pale brown intensity, moderate (++) for intermediate brown intensity, and strong (+++) for intense, dark brown immunoprecipitate. For normal tissues, the percentage of reactivity was individually evaluated for each cell type and calculated as: [(number of DAB-positive cells/total number of cells of the same type) x 100]. Focal staining refers to brachyury staining in less than or equal to 5% of cells.

Immunohistochemical Detection of HER2 and Scoring Method

Formalin-fixed, paraffin-embedded tumor sections were used for immunohistochemical detection of the human epidermal growth factor receptor (HER2) protein. Breast tumor sections were pre-treated using a commercially available kit (Poseidon, Kreatech Diagnostics, Amsterdam) according to the manufacturer's instructions. All tumors were immunostained with the CB11 anti-HER2 Ab (Ventana Medical Systems, Inc. Tucson, AZ). Staining was performed on the Ventana BenchMark XT automated staining platform with the UltraView Universal DAB Detection Kit according to the manufacturer's instructions. HER2 staining in surgical specimens was scored by the pathologist according to the College of American Pathologists (CAP) and the American Society of Clinical Oncology (ASCO) joint guideline on testing for HER2 status in invasive breast cancer (2). Briefly, a score of 0 (negative) was assigned for cases with no discernible membranous staining; 1+ (negative) for weak and/or incomplete membranous staining; 2+ (dubious) for weak to moderate continuous membranous staining in at least 30% of

cells; 3+ (positive) for cases with strong membranous staining in at least 30% of cells. For cases with 2+ staining, the HER2 gene status was detected by fluorescence in situ hybridization using a dual-color probe system that combines a red-labeled HER2 probe (Kreatech Diagnostics) and a green-labeled chromosome 17 centromere probe (CEP17; Kreatech Diagnostics) for enumeration of the ratio of the HER2 gene to chromosome 17. Nuclei were counterstained with DAPI (4',6-diamidino-2-phenylindole, Kreatech Diagnostics). Hybridization analysis was performed as previously described (3). At least 100 tumor cells of the invasive component were analyzed for each case and counts for each probe (HER2 and CEP17) were recorded for each cell/nucleus. A HER2/CEP17 ratio of 2.2 was considered as the cutoff for gene amplification.

Cell Culture

The human breast cancer cell lines used in this study, MCF7 and MDA-MB-436, were obtained from the American Type Culture Collection (ATCC, Gaithersburg, MD) and propagated as recommended. MCF7 cells stably overexpressing human brachyury have previously been described (4). MDA-MB-436 cells were stably transfected with control non-targeting shRNA (designated con.shRNA) or a pool of brachyury-specific shRNA-encoding vectors (designated Br.shRNA) (Sigma-Aldrich) using the Nucleofector technology (Lonza); cells were selected using 1μg/ml puromycin (Sigma-Aldrich).

Real-time PCR

For expression in breast cancer cell lines, total RNA prepared by using the RNeasy kit (Qiagen) was reverse transcribed with the Advantage RT-for-PCR kit (Clontech). For expression in human breast cancer tissues, commercially available normal and tumor breast tissue cDNA panels were

utilized; comprehensive information was reported for the breast tumor tissue samples by the company, including patient staging, patient lymph node status, tumor histological grade, and tumor estrogen receptor (ER), progesterone receptor (PR), and HER2 expression (TissueScan Breast Cancer qPCR Arrays I, II, and III; Origene Technologies, Inc.). For expression in both cell lines and tissues, cDNA (2-100 ng) was amplified using Gene Expression Master Mix and the following Taqman Gene Expression Assays (Applied Biosystems): brachyury (Hs00610080), NANOG (Hs02387400), sex determining region Y box 2 (SOX2) (Hs01053049), octamer-binding transcription factor 4 (OCT4) (Hs00999632), and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (4326317E). Polymerase chain reaction (PCR) was performed according to the manufacturer's recommendations on the 7300 Real-Time PCR System (Applied Biosystems). Mean Ct values for target genes were normalized to mean Ct values for the endogenous control GAPDH ($-\Delta Ct = Ct(GAPDH) - Ct$ (target gene)). The ratio of messenger RNA (mRNA) expression of target gene vs. *GAPDH* was defined as $2^{-\Delta Ct}$.

Brachyury Expression and Prognosis

Breast tumor gene expression data (n=4010) derived from 23 datasets on the Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI) was compiled, including subsets of samples for which recurrence-free survival (n = 1372) and distant metastasis-free survival (n = 2187) were available (5). Each of the 4010 samples was assigned into low (lowest 25%), intermediate, and high (highest 25%) subgroups according to the levels of brachyury mRNA expression (probe set ID: 206524_at). A particular subset of 357 breast cancer patients treated with tamoxifen monotherapy as adjuvant therapy for 5 years post-diagnosis was subsequently chosen for further analysis. By using Kaplan-Meier estimates of survival with the

357 patients, prognosis differences among the three subgroups (low, intermediate, high) were compared. It is important to point out that, by performing the sample classification based on the expression of brachyury in the cohort of 4010 tumors, the sample size in the low/high subgroups corresponding to the 357 patients was not equal. To determine whether brachyury mRNA expression was an independent prognostic factor, a Cox Proportional-Hazards Regression (COXPH) survival analysis was conducted to quantify the weight of the hazard ratios associated with high expression of brachyury and their significance when considered alongside other clinical variables such as tumor size, grade, nodal status, age, HER2, ER and PR status. Statistical analyses were performed using R Project for Statistical Computing (Augasse, Austria), and Kaplan-Meier survival analyses on selected genes were conducted using GraphPad (La Jolla, CA).

Western Blot

Tumor cells were lysed in RIPA (Cell Signaling Technology, Beverly, MA); proteins (20 μg) were resolved using 4-20 % SDS-PAGE, followed by transfer to nitrocellulose membranes. The following antibodies were used: brachyury (ab57480, Abcam, Cambridge, MA) and β-actin (Santa Cruz Biotechnology, Santa Cruz, CA). Blots were incubated with IRDye secondary antibodies (LI-COR Biotechnology, Lincoln, NE); detection was performed with the Odyssey Infrared Imaging System (LI-COR Biotechnology).

Invasion Assays

Blind Well Chambers (Neuro Probe Inc., Gaithersburg, MD) with 10 μm pore polycarbonate filters coated with Matrigel (BD Biosciences, San Jose, CA) were used. RPMI 1640

supplemented with 10% fetal bovine serum (FBS) was added to the lower chambers; cells ($4x10^5$) were added in serum-free medium onto the upper chambers. After 24 hours incubation at 37°C, filters were fixed and stained with Diff-Quik. Cells on the bottom side of the filters were counted in five random X100 microscope objective fields. Experiments were conducted in triplicate for each cell line.

Mammosphere Culture

Methocult H4100 "Base" Methylcellulose Medium (Stem Cell Technologies, Inc.) was mixed at a ratio of 2:3 with Iscove's Modified Dulbecco's Medium (IMDM; Mediatech, Inc.) containing 4 mM glutamine and 1x penicillin/streptomycin (Mediatech, Inc.); epidermal growth factor (EGF), fibroblast growth factor (FGF), insulin, and hydrocortisone were added to attain final concentrations of 20 ng/ml EGF, 20 ng/ml FGF, 5 μg/ml insulin, and 0.5 μg/ml hydrocortisone. Cells were plated in duplicates at 1×10^4 and 5×10^3 cells per well on 24-well ultra-low-attachment plates (Corning, Inc.) for primary and secondary mammosphere culture, respectively. The number of mammospheres per well was counted under 10X magnification following a week of primary or secondary culture.

Tumor Cell Proliferation and Chemotherapy Survival Assays

For proliferation assays, tumor cells were seeded (1×10^3 cells per well) in a 96-well tray, allowed to attach overnight, and incubated for indicated times; MTT reagent (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) (Sigma-Aldrich) was added per well at a final concentration of 0.5mg/ml for 2 hours at 37°C. Medium was removed and the converted dye was solubilized in ice-cold isopropanol. Converted dye was measured at 560nm on a micro

plate reader (Bio-TEK Instruments). Fold growth was calculated relative to the Day-0 value for each cell line. For survival assays, tumor cells were seeded as indicated above, allowed to attach overnight, and treated with indicated doses of docetaxel (Sanofi-Aventis, Bridgewater, NJ) the following day (Day 0). All wells were replaced with fresh media after 6 hours. Cultures were maintained for 5 days, after which cell survival was evaluated by MTT assay as previously described (4). Survival for treated wells was calculated as a percentage of the values representing wells of untreated cells.

Cytotoxicity Assay

Brachyury-specific cytotoxic T lymphocytes (CTL) were generated from the blood of a prostate cancer patient post-vaccination with a PSA-TRICOM-based vaccine (6) as previously described (7). Target cells for cytotoxic assay were pulsed with 50 μCi of ^{111}indium-labeled oxyquinoline. $CD8^+$ T cells negatively isolated from T-cell cultures were used as effector cells at the indicated effector-to-target (E:T) ratios. Following a 16-hour incubation, supernatants were harvested, and released ^{111}In was measured by gamma counting. Spontaneous release was determined by incubating target cells in medium alone, and complete lysis was determined by incubating target cells with 2.5% Triton X-100. All determinations were done in triplicate, and standard error was calculated. Specific lysis was calculated as follows: specific lysis (%) = [(observed release) – (spontaneous release)/(complete release) – (spontaneous release)] x 100.

References

1. Roselli M, Fernando RI, Guadagni F, *et al.* Brachyury, a driver of the epithelial-mesenchymal transition, is overexpressed in human lung tumors: an opportunity for novel interventions against lung cancer. *Clin Cancer Res.* 2012;18(14):3868-79.

2. *HER2 Testing Guidelines and Resources.* http://www.cap.org/apps/cap.portal.

3. Selvarajan S, Bay BH, Mamat SB, *et al.* Detection of HER2/neu gene amplification in archival paraffin-embedded breast cancer tissues by fluorescence in situ hybridization. *Histochem Cell Biol.* 2003;120(3):251-5.

4. Fernando RI, Litzinger M, Trono P, *et al.* The T-box transcription factor Brachyury promotes epithelial-mesenchymal transition in human tumor cells. *J Clin Invest.* 2010;120(2):533-44.

5. Cheng Q, Chang JT, Geradts J, *et al.* Amplification and high-level expression of heat shock protein 90 marks aggressive phenotypes of human epidermal growth factor receptor 2 negative breast cancer. *Breast Cancer Res.* 2012;14:R62

6. Madan RA, Mohebtash M, Arlen PM, *et al.* Ipilimumab and a poxviral vaccine targeting prostate-specific antigen in metastatic castration-resistant prostate cancer: a phase 1 dose-escalation trial. *Lancet Oncol.* 2012;13(5):501-8.

7. Palena C, Polev DE, Tsang KY, *et al.* The human T-box mesodermal transcription factor Brachyury is a candidate target for T-cell-mediated cancer immunotherapy. *Clin Cancer Res.* 2007;13(8):2471-8.

Supplementary Table 1. Brachyury protein expression in paraffin-embedded, formalin-fixed primary breast carcinomas by immunohistochemistry with a murine monoclonal anti-brachyury monoclonal antibody (Abcam, ab57480).

Primary tumor tissue	Brachyury positive (%)
Lymph node status:	
Node-negative	13/15 (86.7%)
Node-positive	14/15 (93.3%)
Grade:	
G1	2/3 (66.7%)
G2	—
G3	25/27 (92.6%)
ER/PR expression:	
ER+ PR+	4/6 (66.7%)
ER- PR-	21/22 (95.5%)
ER/PR/HER2 expression:	
ER+ PR+ HER2+	1/2 (50.0%)
ER+ PR+ HER2-	3/4 (75.0%)
ER- PR- HER2+	12/12 (100.0%)
ER- PR- HER2-	9/10 (90.0%)

G: tumor grade; ER: estrogen receptor; PR: progesterone receptor, HER2: human epidermal growth factor receptor 2

Supplementary Table 2. Clinical characteristics of the patients in each subgroup

Group	*High*	*Intermediate*	*Low*
Age range	44-86	40-88	49-86
Grade 1	16.4%	21.4%	31.0%
Grade 2	60.0%	56.1%	33.3%
Grade 3	23.6%	22.5%	35.7%

Age range for patients included in each of the groups and percentage of samples classified as Grades 1-3 in the high, intermediate and low brachyury mRNA expression groups.

Supplementary Table 3. Cox Proportional-Hazards Regression survival analysis

COXPH analysis (n=270)	*Recurrence-free survival*		*Distant metastasis-free survival*	
	p-value	Coefficient	p-value	Coefficient
Brachyury	0.0027	0.5714	0.0025	0.6375
HER2	0.0436	0.6766	0.0414	0.7354
ER	0.6595	-0.1661	0.7933	0.1142
PR	0.0688	-0.4593	0.0360	-0.5928
Nodal	0.1187	0.3896	0.0703	0.5172
Grade	0.3020	0.1923	0.1741	0.2835
Tumor size	0.0386	0.2135	0.0280	0.2389
Age	0.7937	0.0035	0.7740	0.0044

HER2: human epidermal growth factor receptor 2; ER: estrogen receptor; PR: progesterone receptor. All statistical tests were two-sided.