Supplementary Experimental Procedures

Immunohistochemistry. Staining was performed on OCT or paraffin-embedded hyperplastic mammary glands or breast tumors as previously described (1). All sections were blocked with the Power Block Universal Blocking Agent (Biogenex) according to manufacturer's instructions. Paraffin-embedded sections were subjected to antigen retrieval in a pressure cooker with a sodium citrate buffer and incubated overnight at 4°C with antibodies specific for Ki67 (1:1000; Cat # ab15580, Abcam), F4/80 (1:100; Cat# CL8940AP, Cedarlane), Granzyme B (1:100; Cat# ab4059, Abcam), CD3e (1:200; Cat# ab16669, Abcam), CD8 (1:100; Cat # ab52854, Abcam) and CD20 (1:100; Cat # ab78237, Abcam). OCT-embedded sections were incubated overnight at 4°C with antibodies specific for CD31 (1:200; Cat# 550274, BD Biosciences) or CD4 (1:250; Cat# ab64145, Abcam). All slides were subsequently processed with Vectastain ABC kits (Vector). Paraffin-embedded sections were also subjected to TUNEL staining (Apoptag Detection Kit, Chemicon) according to manufacturer's instructions. Matrigel plugs were injected and stained as previously described (2). Slides were scanned using a ScanScope XT Digital Slide Scanner (Aperio) and quantified with Imagescope software using either the positive pixel count or IHC nuclear algorithms. For immunohistofluorescence, the slides were processed and incubated with primary antibody as described above using the following antibody combinations: CD3e (1:200; Cat# ab16669, Abcam)/Ki67 (1:100; Cat# M7249, Dako) and CD4 (1:200; Cat# ab64145, Abcam)/Ki67 (1:500; Cat# ab15580, Abcam). The appropriate AlexaFluor-488 and AlexaFluor-555 conjugated secondary antibodies were used at a 1:500 dilution (Molecular Probes). Slides were mounted and visualized using an Axiovision 200 microscope (Zeiss).

Tissue microarray. The cohort consisted of 179 invasive breast cancers diagnosed at the Department of Pathology, Malmo University Hospital, Sweden, between 2001 and 2002. The

median age at diagnosis was 65 (range 35-97) and follow-up period for overall survival was 52 months (range 4-63). The study was approved by the ethical committee at Lund University. Tissue microarrays were constructed using a manual tissue arrayer (Beecher Inc.). Two 1.0 mm cores from each donor block were assembled into recipient blocks. IHC analysis was performed on 4 μm sections in the Ventana Benchmark system using pre-diluted antibodies to ER, PR and HER2. For CD8 detection, antigen retrieval was performed using microwave treatment for 7 minutes in a citrate buffer (BioGenex). For ShcA detection, antigen retrieval was performed using a pressure cooker for 10 minutes in a 10 mM TE pH9 buffer. Endogenous peroxidase activity was blocked by incubating slides in 3% hydrogen peroxide (Sigma) for 20 minutes. To reduce nonspecific background, slides were pretreated with blocking buffer containing either 10% goat serum/1% BSA (CD8) or Power Block (Biogenex - ShcA) for 30 minutes. Primary antibodies were pre-diluted in blocking buffer (CD8-1:100; C8/144B, NeoMarkers) or 3% BSA (ShcA-1:2500; sc-288, Santa Cruz) and applied to tissue section for 16 hours at 4°C. Slides were developed with either the Ultravision LP Detection System (Thermo Scientific-CD8) or Envision Plus HRO Rabbit Dab Plus System (Dako - ShcA). 3'-diaminobenzidine (DAB) was used as a chromogenic substrate and slides were counterstained using haematoxylin.

Digital images were captured using the Aperio ScanScope XT Slide Scanner (Aperio Technologies). Slides were analyzed using positive pixel count or IHC nuclear algorithms.For the TMA, ShcA was analyzed using the positive pixel count algorithm and CD8, was analyzed using Genie pattern recognition software and IHC nuclear algorithm. For the TMA, calculations were performed using SPSS version 15.0 (SPSS Inc, Chicago). For each patient, the ShcA autoscore was calculated by multiplying the average staining intensity with the percentage of positively stained cells over duplicate cores. For each core, the density of the CD8 cells (number

of cells/mm^2) was quantified and the cut off for analysis was generated using decision tree analysis.

FACS Analysis. Tumors were finely cut and digested in DMEM cell culture medium containing Liberase blendzyme 3 (0.9 Wunsch Units/ml) (Roche), DNaseI (200 μg/ml) (Roche) and penicillin (100 Units/ml)/streptomycin (100 μg/ml) (Gibco/Invitrogen) for 30 minutes at 37°C with gentle agitation. Cell suspensions were pelleted and further digested in a non-enzymatic tissue dissociation buffer (Gibco/Invitrogen) at room temperature for 10 minutes. Remaining tissue debris was removed by filtration, and red blood cells were removed by treatment with Ack-lysis buffer. Cells were resuspended in PBS and stained with CD4-PECy7 (RM4-5) and Foxp3-PE (FJK-16s) (eBioscience) antibodies. The samples were analyzed on a FACSCalibur (BD Biosciences) using the FlowJo software.

ELISAs. Lysates from flash frozen tumor tissue were generated as previously described (2) and 75-100 μg total protein was subjected to ELISA analysis using the following mouse kits: CCL2 (Cat # MJE00; R&D) , CCL3 (Cat # MMA00; R&D), CCL5 (Cat# MMR00; R&D), CXCL9 (Cat# MCX900; R&D), CXCL10 (Cat# MCX100; R&D), IL2 (Cat# 88-7024-86; Ebioscience), IL4 (Cat# 88-7044-86; Ebioscience), IL10 (Cat# 88-7104-86; Ebioscience), IFNγ (Cat #88-8314-86; Ebioscience) and VEGF (Cat # MCMV00; R&D). Equivalent amounts of protein were normalized to total tubulin levels by fluorescent immunoblot analysis and visualized with an Odyssey infrared scanner (Li-COR Biosciences).

Immunoblotting. Lysates from flash frozen tumor tissue were generated as previously described (2) and 15 μg total protein was subjected to immunoblot analysis using the following antibodies: ShcA (1:2500; Cat# S14630 - BD Biosciences), ShcB (1:1000; Cat# sc-33808 – Santa Cruz), ShcC (1:1000; Cat# sc-28833 - Santa Cruz), ShcD (1:500 – generous gift from Heike

Wegmeyer/Tony Pawson), ErbB2 (1:2500; Cat# sc-284 - Santa Cruz), Cre (1:1000; Cat# PRB106C - Covance), anti-mouse-IgG (1:10000; Cat# 715-065-150 – Jackson Labs), anti-mouse IgM, IgG1, IgG2a, IgG2b, IgG3 (1:10000; Cat# 715-065-020, 115-065-205, 115-065-206, 115-065-206, 115-065-209 – Jackson Labs), pAKT (1:1000; Cat# 9271 - Cell Signaling), AKT (1:1000; Cat# 9272 - Cell Signaling), pERK (1:1000; Cat# 9106 - Cell Signaling), ERK (1:1000; Cat# 9102 - Cell Signaling), pp38-MAPK (1:1000; Cat# 9215 - Cell Signaling), p38-MAPK (1:1000; Cat# 9212 - Cell Signaling), pJNK (1:1000; Cat# 9252 - Cell Signaling), JNK (1:1000; Cat# 9251 - Cell Signaling), β-actin (1:5000; Cat #AB5441 – Sigma) and α-Tubulin (1:5000; Cat# 2125 – Cell Signaling). The blots were incubated with HRP-conjugated secondary antibodies or streptavidin-HRP (Jackson Labs) and visualized by ECL (Amersham). Fluorescent immunoblotting was performed using the appropriate fluorescently-labeled secondary antibodies and the Odyssey infrared scanner (Li-COR Biosciences). Relative IgG levels were calculated by dividing the signal obtained with an anti-mouse IgG antibody from the signal obtained with an anti-tubulin antibody.

Quantitative real-time reverse transcription-PCR. Total RNA was isolated from tumor tissue using RNeasy midi-kits (Qiagen) according to manufacturer's instruction. For quantitative real RT-PCR, total RNA was used with the QuantiTect SYBR Green RT-PCR kit (Qiagen) and the LightCycler (Roche). A constant annealing temperature of 56°C was used for all reactions. Each amplification reaction was performed in triplicate and transcript levels were normalized to GAPDH. In Figure S7, CD28 and ICOS transcript levels were also normalized to the fold change in CXCR3 expression levels between individual tumors. Primer sequences, amount of input RNA and the detection temperature are described in Supplementary Table 6.

Statistical Analysis Unless indicated, all statistical analyses were performed using a student's t test. Statistical analysis for the proliferating CD4+ cells (Fig. 3) used a Fisher's exact test. For Fig. 5C, 5D 6B and 6C, statistical analysis was performed using the χ^2 test. The Kaplan-Meier analysis and the log rank test illustrated differences between survival and ShcA expression (Figs. 5B, 7). All tests were two-tailed.

Supplementary Figure Legends

Supplementary Figure 1: The remaining Shc family members are not overexpressed in NIC/ShcA$^{fl/fl}$ mammary tumors. **(A)** Relative ShcA, ShcB, ShcC and ShcD transcript levels, normalized to GAPDH, in NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors (n=12 for each genotype). **(B)** Immunoblot analysis on tumor cell lysates from the indicated genotypes using ShcB, ShcC and actin-specific antibodies. **(C)** Tumor lysates from the indicated genotypes were immunoprecipitated then immunoblotted with a ShcD-specific antibody. Numbers identify individual mice.

Supplementary Figure 2: The lung metastatic burden is similar in NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ transgenic mice. Lung metastases were scored by examining H&E-stained sections of all five lung lobes. Five, 50 micron step sections were analyzed for each animal. **(A)** Percentage of animals with lung metastases. N represents the number of animals analyzed for each genotype. **(B)** The number of metastatic lesions observed within all five lung lobes for the genotypes indicated. N represents the number of animals analyzed. **(C)** The total tumor burden within the lungs was determined by calculating the total lung area containing metastatic lesions (mm^2). N represents the number of animals analyzed.

Supplementary Figure 3: NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ breast tumors display a similar proliferative, apoptotic and angiogenic potential. **(A)** Immunohistochemical staining on paraffin-embedded mammary tumors using a Ki67-specific antibody for the indicated genotypes. The data is representative of 6 independent mammary tumors (10-20 fields/tumor; 20x magnification) and is depicted as % Ki67 positive cells ± SEM. **(B)** TUNEL staining on paraffin-embedded tumors for the indicated genotypes. The data is representative of 5 independent tumors (10-20 fields/tumor; 20x magnification) and is depicted as % TUNEL positive cells ± SEM. **(C)** Immunohistochemical staining of OCT-embedded tumors with a CD31-specific antibody for the indicated genotypes. The data is representative of 5 independent tumors (10-20 fields/tumor; 20x magnification) and is depicted as % CD31 positive area/total epithelial area ± SEM. **(D)** Relative VEGF protein levels, as determined by ELISA, in whole cell lysates from NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. **(E)** Three NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ primary breast cancer cells were disaggregated and subjected to Matrigel plugs assays. OCT-embedded Matrigel plugs were subjected to CD31 immunohistochemical staining and the number of CD31 positive pixels per nucleus in each plug was quantified. The data is representative of 3 independent Matrigel plugs for each individual tumor. Numbers identify individual mice.

Supplementary Figure 4: The MAPK and AKT signaling pathways are not significantly altered in NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. Immunoblot analysis on tumor cell lysates from the indicated genotypes using phospho-ERK, ERK, phospho-p38 MAPK, p38 MAPK, phospho-JNK, JNK, phospho-AKT and AKT specific antibodies. Numbers identify individual mice.

Supplementary Figure 5: Macrophage infiltration is similar in NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. **(A)** Relative CCL2, CCL3 and CCL5 protein levels, as determined by ELISA,

in whole cell lysates from NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. **(B)** Immunohistochemical staining on paraffin-embedded mammary tumors using a F4/80-specific antibody for the indicated genotypes. The data is representative of 6 independent mammary tumors (10-20 fields/tumor; 20x magnification) and is depicted as % F4/80 positive cells ± SEM.

Supplementary Figure 6: Increased CXCL9 expression in pre-neoplastic mammary glands from MT-Y250F mice. Macrophage infiltration is similar in NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. Relative CXCL9, CCL2 and CCL3 protein levels, as determined by ELISA, in whole cell lysates (n=7 for each) from MT and MT-Y250F mammary tumors and adjacent mammary tissue (AMG), which predominately contains pre-neoplastic structures. P values associated with changes in chemokine levels between MT and MT-Y250F lysates from AMG* or BT** tissue are as follows: CXCL9 (AMG-0.09; BT-0.71), CCL2 (AMG-0.09; BT-0.07), CCL3 (AMG-0.69; BT-0.4). CXCL10 and CCL5 protein levels were undetectable in any of the tissue examined.

Supplementary Figure 7: Increased T cell activation in NIC/ShcA$^{fl/fl}$ mammary tumors. **(A)** Relative CXCR3, CD28 and ICOS transcript levels, normalized to GAPDH, in NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. The relative CD28 and ICOS expression levels were also normalized to the relative fold change in CXCR3 levels within individual tumors. **(B)** Relative IL2 protein levels, as determined by ELISA, in whole cell lysates from NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. **(C)** Immunofluorescent staining of paraffin-embedded NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors with CD3e (green) and Ki67 (red) specific antibodies. White arrows indicate proliferating CD3e+ cells. The images were taken at 63X magnification.

Supplementary Figure 8: Infiltration of CD8 and Granzyme B positive cells is exceedingly low in mammary tumors irrespective of their ShcA status. Immunohistochemical staining on paraffin-embedded mammary tumors using a **(A)** Granzyme B or **(B, C)** CD8-specific antibody for the indicated genotypes. The data is representative of 5-6 independent mammary tumors (10-20 fields/tumor; 20x magnification) and is depicted as % positive cells ± SEM.

Supplementary Figure 9: Cytokine levels are similar in end-stage NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. **(A)** Percentage of CD4+/FoxP3+ T cells between NIC and NIC/ShcA$^{fl/fl}$ mammary tumors as assessed by flow cytometry (n=4 for each genotype). **(B)** Relative IFNγ, IL4 and IL10 protein levels, as determined by ELISA, in whole cell lysates from NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors (n=9 for each).

Supplementary Figure 10: Increased immunoglobulin synthesis and class switch recombination in NIC/ShcA$^{fl/fl}$ mammary tumors. **(A)** Relative IgJ, IgHM, IgHG1 and IgHG2a transcript levels, normalized to GAPDH, in NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ mammary tumors. **(B)** Immunoblot analysis on lysates from NIC/ShcA$^{+/+}$ and four NIC/ShcA$^{fl/fl}$ mammary tumors using IgHM, IgHG1, IgHG2a, IgHG2b, IgHG3, Cre, ShcA and β-Actin-specific antibodies. Numbers identify individual mice. **(C)** Immunohistochemical staining of paraffin-embedded NIC/ShcA$^{+/+}$ and NIC/ShcA$^{fl/fl}$ tumors with a CD20-specific antibody. The data is representative of six independent tumors (10-20 fields/tumor at 20x magnification) and is depicted as % CD20^{+} cells ± SEM. Representative images demonstrating CD20 infiltration are depicted.

Supplementary Figure 11: Increased IgG levels in hyperplastic mammary glands and breast tumors from MT-Y250F animals. Immunoblot analysis for IgG and Tubulin levels on the indicated whole cell lysates. Fluorescently-labelled secondary antibodies permitted

quantification of relative IgG levels, normalized to tubulin levels, for each sample. The graph represents the relative IgG protein levels within hyperplastic mammary glands (n=9) and breast tumors (n=9) in MT and MT-Y250F animals.

Supplementary Figure 12: The SRIS only discriminates the stromal contribution of the breast tumor. Heatmaps from two publically available datasets that utilized laser capture microdissection to specifically profile the tumor epithelium versus stroma in primary breast cancers. The SRIS is only able to define groups of patients overexpressing this signature when utilizing stromally-derived gene expression profiles. For each patient, their outcome (red = death) and molecular subtypes (dark blue=luminal A, light blue=luminal B, green=normal, pink=Her2, red=basal) are shown. For the Boersma et al. dataset, the blue bars indicate those patients with inflammatory breast cancer.

Supplementary Figure 13: The SRIS predicts good outcome within the Her2 and basal subtypes even in individual datasets. Kaplan-Meier analysis within the SRIS-Low (blue) and SRIS-High (red) groups for each of the molecular subtypes. The number of patients within each subtype is shown. HR = hazards ratio.

References

1. Ursini-Siegel J, Rajput AB, Lu H, Sanguin-Gendreau V, Zuo D, Papavasiliou V, et al. Elevated expression of DecR1 impairs ErbB2/Neu-induced mammary tumor development. Mol Cell Biol. 2007;27:6361-71.
2. Ursini-Siegel J, Hardy WR, Zuo D, Lam SH, Sanguin-Gendreau V, Cardiff RD, et al. ShcA signalling is essential for tumour progression in mouse models of human breast cancer. EMBO J. 2008;27:910-20.