Supplementary Information

**Supplementary Figure 1.** Immunohistochemical analyzes of HMGA2 expression are shown (40x). These are representative sections from staining performed on patient tumor samples described in Table 1. *Scale bars represent 250 μm.*

**Supplementary Figure 2.** Agarose Colony Formation Assay. Anchorage-independent proliferation was performed by plating SW403, HT29, SW620, SW480 and HCT116 (1000 cells per well) in six-well plates, with a bottom layer of 0.6% agar and a top layer of 0.3% agar containing the cells. After 21 days incubation, colonies were counted.

**Supplementary Figure 3.** Immunofluorescence analysis of ZEB1 (left), Fibronectin (right) expression in MCF7-Mock and MCF7-HMGA2 cells. *Scale bars represent 50 μm.* ZEB1 and fibronectin are expressed in the nucleus of MCF7-HMGA2 cells that have undergone the EMT as is consistent with the other EMT markers in Figure 1C.

**Supplementary Figure 4.** β-catenin was relocalized to the nucleus at the invasive front of Wnt1-mediated tumor on Hmga2^{+/+} genotypic backgrounds, but not on the Hmga2^{−/−} genetic background. β-catenin was detected in the nucleus of the cells at the invasive front consistent with the HMGA2 expressed cells in the Wnt1-mediated tumor on the Hmga2^{+/+} genetic background (arrows). On the
other hand, no β-catenin relocalization was detected in tumor on the Hmga2−/−
genotypic backgrounds. Scale bars represent 100μm.

**Supplementary Figure 5.** TGFβRII was up-regulated in 4T1 cells. TGFβRII mRNA expression was highly up-regulated in 4T1 cells compared to 4TO7 parental cells (## = P < 0.0001).

**Supplementary Figure 6.** (A) HMGA2 was detected at the edge of the metastatic lung tumor (left, arrows) and from the vessel of the liver stroma to the parenchyma in the metastatic liver tumor (right, arrows). Scale bars represent 100 μm. (B) Liver metastases (arrows) were induced in 4TO7 cells stably over-expressing HMGA2 (4TO7-HMGA2).

**Supplementary Figure 7.** Schematic of liver metastasis. Metastatic cancer cells (4TO7 cells stably over-expressing HMGA2 and 4T1 cells) are located in Glisson’s capsule (GC) and Liver parenchyma (LP).

**Supplementary Figure 8.** There is no statistically significant difference in Snail and Twist mRNA expression in ectopic HMGA2 transfected cells (MCF7-HMGA2 cells).

**Supplementary Figure 9.** TGFβ 1 mRNA expression in MCF7, MCF-HMGA2, and MDA-MB231 human breast cancer cells as detected by qRT-PCR.
Supplementary Figure 10. HMGA2 mRNA expression in MCF7 and MDA-MB231 human breast cancer cells treated with TGFβ 1 as detected by qRT-PCR.

Supplementary Materials and Methods

SiRNA transfection.

MDA-MB231 cells were transfected with Hmga2 siRNA using Lipofectamine RNAiMAX (Invitrogen). Lipid complexes were incubated for 10 minutes on the 6-well plate and $3 \times 10^5$ cells per 2.5 ml of complete growth medium without antibiotics were plated onto the complexes. Cells were harvested 12, 24 and 48 hours post transfection for mRNA and protein analysis.

Quantitative Evaluation of Human Samples.

Two independent observers evaluated all human samples microscopically. The proportions (i.e. percent of positive cells per total number of cells) of HMGA2 and Ki67-positive colonic epithelial cells were scored by counting at least 200 continuous epithelial cells in randomly selected fields at both the invasive front and throughout the entire neoplasm using a standardized grid (Olympus microscope BX40). Representative images are shown in Supplementary Fig. 1 delineating the invasive front in HMGA2 stained sections of tumor.

Quantitative Evaluation in Mice.
The mice were examined for a period of 50 weeks at weekly intervals for palpable tumors. After sacrificing the mice, tumors were removed and tissues were stored, either snap-frozen or fixed in 4% paraformaldehyde for histopathology and immunohistochemical analyses. All counting was performed by two independent observers blind to the genotype of the mice.