

## Additional file legend list:

### 1. Additional File 1, Figure S1

Title: Six1 is enriched in the TIC population of luminal tumors

Description:

- (A) Relapse free survival curve of Kaplan-Meier analyses with a combined 779 tumor data set. Kaplan Meier curves are shown for Basal tumors, Claudin low tumors and Her2 positive tumors.
- (B) Flow cytometry was performed on the luminal B patient xenografts PK12, PK15 and PE4 following staining of the cells with antibodies that recognize CD24 and CD44.
- (C and D) The levels of Six1 mRNA are increased in MCF7 (B) and T47D (C) secondary tumorspheres as compared to the parental cells. Error bars represent mean value +/- SEM. *P* values represent statistical analysis using a two-tailed *t* test. The experiments were performed at least 3 times.

### 2. Additional File 2, Figure S2

Title: Six1 overexpression enhances TIC characteristics

Description:

- (A) Six1 differentially regulates genes in the tumor initiating cell signature. The cancer stem cell/tumor initiating cell gene list was obtained from Liu et al<sup>45</sup>. Microarray data from the MCF7-control and MCF7-Six1 expressing clones were filtered for genes included in the cancer stem cell/tumor initiating cell gene list and expression values across the samples were hierarchically clustered. The included gene expression data represents those genes that are consistently regulated across the clones. The color scale represents the expression level of genes above (red), below (green), and at (black) the mean expression level of that gene across all samples.
- (B) Six1 overexpression leads to an increase in tumor initiating cell characteristics. Flow cytometry was performed on MCF7-Six1 and MCF7-Ctrl cells following staining of the cells with antibodies that recognize CD24 and CD44. The boxed region represents the CD24<sup>low</sup> CD44<sup>+</sup> TIC population.
- (C) Six1 overexpression in MCF7 cells promotes tumor initiation in NOD/SCID mice. 10<sup>5</sup>, 10<sup>4</sup> and 10<sup>3</sup> MCF7-Ctrl or MCF7-Six1 cells were injected into the number 4 mammary fat pad of 6-week old female NOD/SCID mice, and the mice were then palpated weekly for tumor formation. Data is shown is from 2 weeks to 4 weeks post injection.

### 3. Additional File 3, Figure S3

Title: Single cell tumorsphere assays demonstrate that Six1 overexpression increases the functional tumor initiating population.

Description:

- (A) Single cell tumorsphere assays demonstrate that Six1 overexpression increases the functional tumor initiating population. Secondary tumorsphere assays were performed by plating single cells from the primary tumorsphere in 96 well ultra low attachment plates and culturing for 10 days. The graph represents 3

individual clones. Error bars represent mean value  $\pm$  SEM. *P* values represent statistical analysis using a two-tailed *t* test.

(B) Representative images of tumorspheres.

#### 4. Additional File 4, Figure S4

Title: TGF- $\beta$  signaling is in part required for the ability of Six1 to induce TICs and pERK signaling

Description:

- (A) SB431542 inhibits TGF- $\beta$  signaling in both MCF7-Ctrl and MCF7-Six1 cells, as assessed by TGF- $\beta$  responsive promoter 3TP-luciferase activity after normalizing to renilla-luciferase. Error bars represent mean value  $\pm$  SEM. *P* values represent statistical analysis using a two-tailed *t* test. The experiments were performed at least 3 times.
- (B) Inhibition of TGF- $\beta$  signaling by SB431542 reverses tumorsphere forming efficiency in MCF7-Six1 cells back to levels observed in the control cells. Representative secondary tumorsphere assay was shown after 10 days of culture. The experiments were performed at least 3 times. Error bars represent mean value  $\pm$  SEM. *P* values represent statistical analysis using a two-tailed *t* test.
- (C) T $\beta$ RIIDN inhibits TGF- $\beta$  signaling in both MCF7-Ctrl and MCF7-Six1 cells, as assessed using the TGF- $\beta$  responsive promoter 3TP-luciferase activity after normalizing to renilla-luciferase. Error bars represent mean value  $\pm$  SEM. *P* values represent statistical analysis using a two-tailed *t* test. The experiments were performed at least 3 times.
- (D) Six1 expands the MCF7 TIC population through activating TGF- $\beta$  signaling. Decreasing numbers of MCF7-Six1, MCF7-Six1/T $\beta$ RIIDN or MCF7-Ctrl cells were injected into the number 4 mammary fat pad of 6-week old female NOD/SCID mice, and the mice were then palpated weekly for tumor formation. The data from 2,3 and 4 weeks post injection of tumor cells are shown.
- (E) Tumor volumes do not differ between MCF7-Six1/GFP and MCF7-Six1/T $\beta$ RIIDN groups. Tumor volumes were measured 5 weeks post injection. The data shown is taken from tumors that formed in the group of mice injected with  $10^3$  cells. The tumor size was not significantly different between the MCF7-Six1/GFP and MCF7-Six1/T $\beta$ RIIDN. *P* values represent statistical analysis using 1 way ANOVA.
- (F) The TGF- $\beta$  type I receptor kinase inhibitor, SB431542, partially, but not completely, represses ERK phosphorylation activated by Six1. MCF7-Ctrl and MCF7-Six1 cells (3 clones each) were treated with 1mM SB431542 or vehicle (DMSO) for 48 hours after which whole cell lysates were collected and Western blot analysis was performed for p-ERK. The quantification is shown in Figure 3 (E).

#### 5. Additional File 5, Figure S5

Title: MEK1/2 inhibitors inhibit EMT and TIC characteristics

Description:

- (A) The MEK1/2 inhibitor U0126 decreases pERK in MCF7-Six1 and MCF7-Ctrl cells. MCF7-Ctrl and MCF7-Six1 (3 individual clones each) were treated with 10mM U0126 or vehicle (DMSO) for 2 hours. The whole cell lysates were used to perform western blot analysis using anti-pERK (1:1000) and anti-total ERK (1:1000) antibodies.
- (B and C) U0126 inhibits the relocalization of E-cadherin (B) and  $\beta$ -catenin (C) observed with Six1 overexpression. Graphs represent quantitation of Western blots examining the soluble vs. insoluble E-cadherin and  $\beta$ -catenin. The data represent mean value  $\pm$  SEM. *P* values represent statistical analysis using a two-tailed *t* test.
- (D) The MEK1/2 inhibitor AZD6244 decreases pERK in MCF7-Six1 and MCF7-Ctrl cells. MCF7-Ctrl and MCF7-Six1 (3 individual clones each) cells were treated with 10mM AZD6244 or vehicle (DMSO) for 2 hours. Whole cell lysates were extracted and used to perform Western blot analysis with anti-pERK (1:1000) and anti-total ERK (1:1000) antibodies.
- (E) Inhibition of MEK1/2 with AZD6244 dramatically reduces the secondary tumorsphere formation of MCF7-Six1 cells. Representative secondary tumorsphere assay is shown after 10 days of culture. The experiments were performed at least 3 times. Error bars represent mean value  $\pm$  SEM. *P* values represent statistical analysis using a two-tailed *t* test.
- (F) Inhibition of MEK1/2 by AZD6244 dramatically decreases tumor formation efficiency in NOD/SCID mice.  $10^6$  MCF7-Six1 or MCF7-Ctrl cells (3 individual clones each) were injected underneath the nipple of the 4th mammary fat pad of 6-week old female NOD/SCID mice. One week post injection, mice were treated with AZD6244 by oral gavage twice per day for 3 days and once per day for the following 3 days (vehicle, 25mg/kg or 50mg/kg). The mice were palpated weekly for tumor formation. The data from 2,3 and 4 weeks post injection of tumor cells are shown.

## 6. Additional File 6, Figure S6

Title: Six1 KD inhibits both tumor initiation and metastasis in an immunocompetent model, and the MEK1/2 inhibitor, AZD6244, inhibits primary tumor burden similarly to Six1 KD

Description:

- (A) Knockdown of Six1 inhibits tumor initiation. 66cl4 scramble control or Six1KD cells were serially diluted and then injected into the 4th mammary fat pad of 6-week old female BALB/c mice. Mice were monitored for tumor initiation weekly. The data from 2,3, and 4 weeks post injection of tumor cells are shown.
- (B)  $10^6$  66cl4/scramble or 66cl4/Six1KD cells were injected into the 4th mammary fat pad. One week post injection, mice were treated by oral gavage with vehicle or AZD6244 at a concentration of 50mg/kg or vehicle twice per day for 7 days. After injection of 150mg of luciferin/kg into the mice, IVIS imaging was used to quantitate the primary burden at 3 weeks post injection. Error bars represent mean value  $\pm$  SEM. It should be noted that the signal from a number (5/10) of the scramble control tumors was saturated at this time point, suggesting that the size

of these tumors may be underestimated. In contrast, only 1/8 of 66cl4/scramble tumors in the AZD6244 group had saturated signal, and none of the tumors in the 66cl4/Six1 KD group had saturated signals.

#### **References:**

45. Liu, R. *et al.* The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* **356**, 217-226, doi:356/3/217 [pii] 10.1056/NEJMoa063994 (2007).