(Tam-R) was immunoblotted. **b.** ChIP of ER, Pax2, AIB-1 and HDAC-1 at the ER binding site in the *ERBB2* gene in wild type and Tam-R cells after vehicle or tamoxifen treatment. **c.** Control or Pax2 expressing plasmids were transfected into Tam-R cells, followed by vehicle (V) or tamoxifen (T) treatment. Total protein was immunoblotted. Following Pax2 expression in Tam-R cells, total viable cell numbers were determined after vehicle or tamoxifen treatment. The immunoblots were cropped and the original figures are in Supplementary data 11. All experiments are the average of three independent replicates ± Std Dev.

**Figure 4.** Pax2 predicts clinical outcome in breast cancer patients. Kaplan-Meier curve representing percent relapse free survival in tumours based on Pax2 and AIB-1 staining (n = 109). The percentage of ErbB2 over-expressing tumours within each category is provided.

**Supplementary Methods**

Cell lines

MCF-7 cells, T47D, ZR75-1 and BT-474 cells were grown as previously described[27]. Tam-R cells were derived by long term exposure to tamoxifen[18] and grown in the same conditions as wild type MCF-7 cells. Herceptin was added to the media at a final concentration of 10μM.

ChIP-on-chip Experiments

Genome-wide Estrogen Receptor ChIP-on-chip experiments were performed in duplicate, as previously described[2], with the exception that the Affymetrix seven Genechip tiling array 2.0R set (Catalogue number 900772) were used. Analysis of ER binding sites were determined using MAT[28], with a False Discovery Rate (FDR) of 5%. Microarray data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-TABM-563.

Motif enrichment

Analysis of motif enrichments were performed using the CEAS program (http://ceas.cbi.pku.edu.cn/). The Pax motif is represented using Weblogo (http://weblogo.berkeley.edu/).

Plasmids

Pax2 expression was from p-TARGET-Pax2 (a kind gift from Dr. Stefano Buttiglieri, Torino, Italy), AIB-1 expression was from pcDNA-AIB-1 construct (a kind gift from Dr. Jerome Eeckhoute, France) and SRC-1 expression was from pSG5-SRC-1.

RT-PCR

Cells were deprived of hormones as previously described[29]. Total RNA was collected and RT-PCR was performed as previously described[2]. Primer sequences are provided in Supplementary data 15.

Chromatin Immunoprecipitation (ChIP)

ChIP experiments were performed as previously described[3]. Antibodies used were: ERα (HC-20), AIB-1/RAC3 (M-397) and HDAC-1 (sc-6298 and sc-6299) from Santa Cruz (Santa Cruz Biotechnologies, CA), RNA PolII (ab5408), H3R17 dimethyl (ab8284), Pax2 (ab23799), SRC-1 (ab2859) and N-CoR (ab24552) from Abcam, UK. Primer sequences are provided in Supplementary data 15.

siRNA

siRNA experiments were as previously described[2]. The sequence of the siRNAs were: siPax2 (Sequence 1): sense GAAGUCAAGUCGAGUCUAUUU and antisense: AUAGACUCGACUUGACUUCUU; siPax2 (Sequence 2): sense CAUCAGAGCACAUCAAAUCUU and antisense: GAUUUGAUGUGCUCUGAUGUU (Dharmacon, USA); siN-CoR Smartpool (Cat. Number M-003518-01, Dharmacon, USA), containing the following sequences: GAUCACAUCUGUCAAAUUAUU, GAACGUGGCUCUCAAAGUUU, GAAAGGAAAUCGACACUGAUU and GCCCUGGGAUUUAUGAUGAUU.

Western blotting

Cells were deprived of hormones as previously described[29]. Antibodies used were: ERα (Ab-10) from Neomarkers (Lab Vision, UK); ErbB2 (ab16901), Pax2 (ab38738), SRC-1 (ab2859), N-CoR (ab24552) and β-Actin (ab6276) from Abcam, UK; AIB-1/RAC3 (M-397) from Santa Cruz (Santa Cruz Biotechnologies, CA).

Cell counting

Cells were plated at equal confluence and grown in full DMEM media. Cells were transfected as previously described[2] and cells were stimulated with 100nM oestrogen or 1μM 4-hydroxy-tamoxifen for 24 hours, or in time course experiments for the time periods given in the figure. Total cells were harvested for automated cell counting using the Z2 Coulter Particle Count Analyzer.

Chromosome Conformation Capture assay (CCC)

CCC was performed according to published protocols[30]. The chromatin was digested with PstI and the real time primers used were Fwd: GGAGCGGAAGTGATTCAGAG; Rev: TTGCAGAGACCTCTGGGAGT. The taqman probe was FAM-AGAGCAGTTCTGCTCTTCGC. A control reverse primer against another PstI site was included; AGAGTCACCAGCCTCTGCAT.

Immunohistochemistry

109 ER positive breast cancer sections were collected and processed as previously described[25]. Immunohistochemistry for Pax2 was performed on an automated BondMax Immunostainer (Leica, Germany) using Pax2 antibody (ab23799, Abcam) at a concentration of 1:100. AIB-1 immunohistochemistry was performed using AIB-1 antibody (611105, Transduction laboratories, UK) at 1:200. Immunohistochemistry for ErbB2 was previously described[25]. Standardisation of Pax2 scoring was achieved by comparison of scores between two independent individuals. Scores were given as the percentage of nuclei staining positive (within multiple representative areas of the tumour) and the absolute intensity of nuclear staining on a scale of 0 to +++ (0, no staining; +, slight/weak heterogenous nuclear staining; ++, strong homogenous nuclear staining; and +++, intense homogenous nuclear staining). Examples of Pax2 positive and negative stained samples are provided in Supplementary data 16.

Statistical analyses

All statistical analyses were performed using two tailed paired T-tests and only values lower than p-value 0.05 were considered statistical. In all figures, the data are the average of a minimum of three independent replicates ± Std Dev. Kaplan-Meier curve statistics were determined using log-rank test. Pax2 and AIB-1 relationships were

established using Cox regression analysis. Analysis was carried out in SPSS V16.0 for Mac (SPSS inc, Chicago, Il). Kaplan-Meier plots were constructed to display the data and analysis was carried out using Cox Regression. Time to relapse was taken as the outcome, binary variables for AIB-1 and Pax2 and the interaction between AIB-1 and Pax2 were included as predictor variables.

## References

[30] H. Hagege, P. Klous, C. Braem et al., Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nature protocols* **2** (7), 1722 (2007).