

Supplementary Information

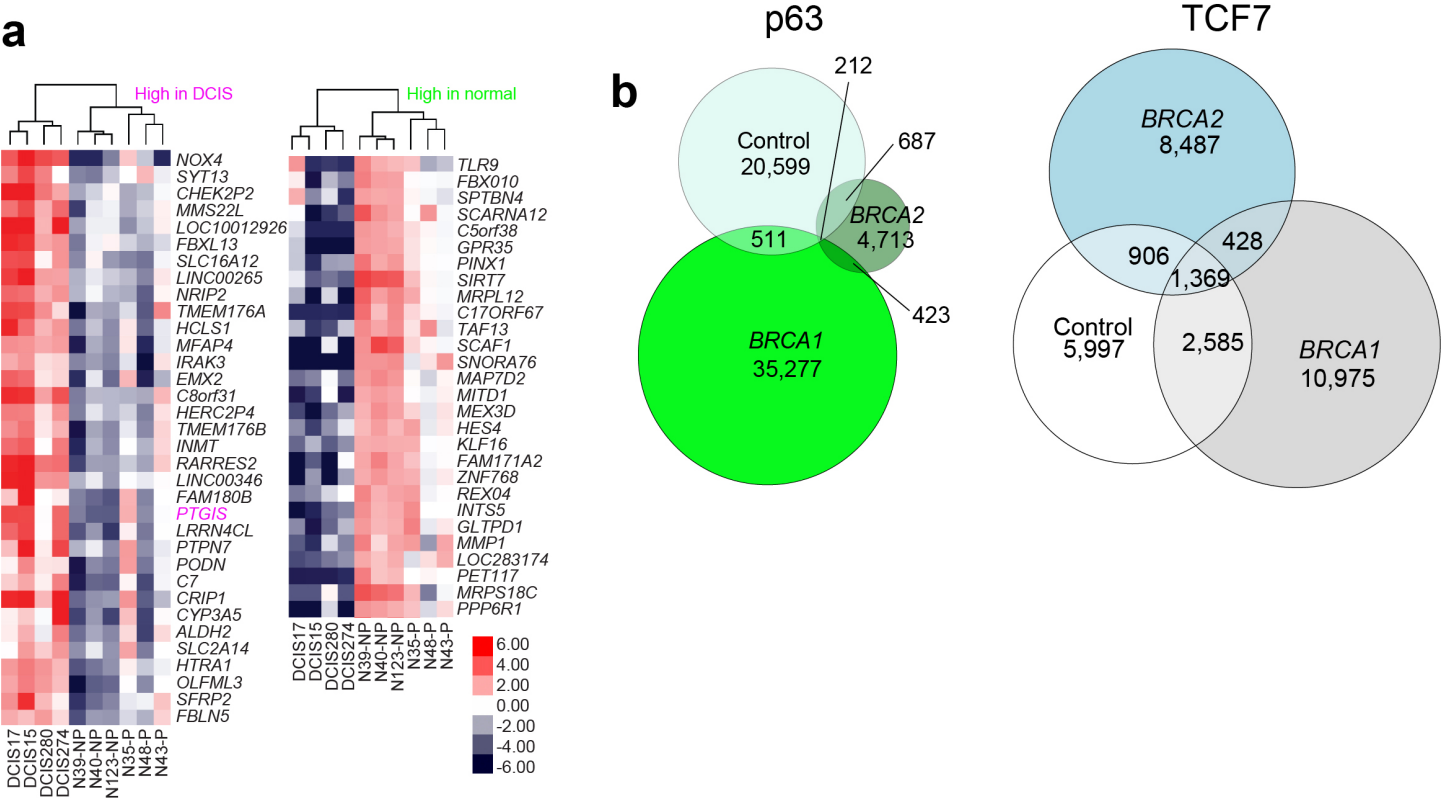
Perturbed myoepithelial cell differentiation in *BRCA* mutation carriers and in ductal carcinoma *in situ* (DCIS)

Ding et al.

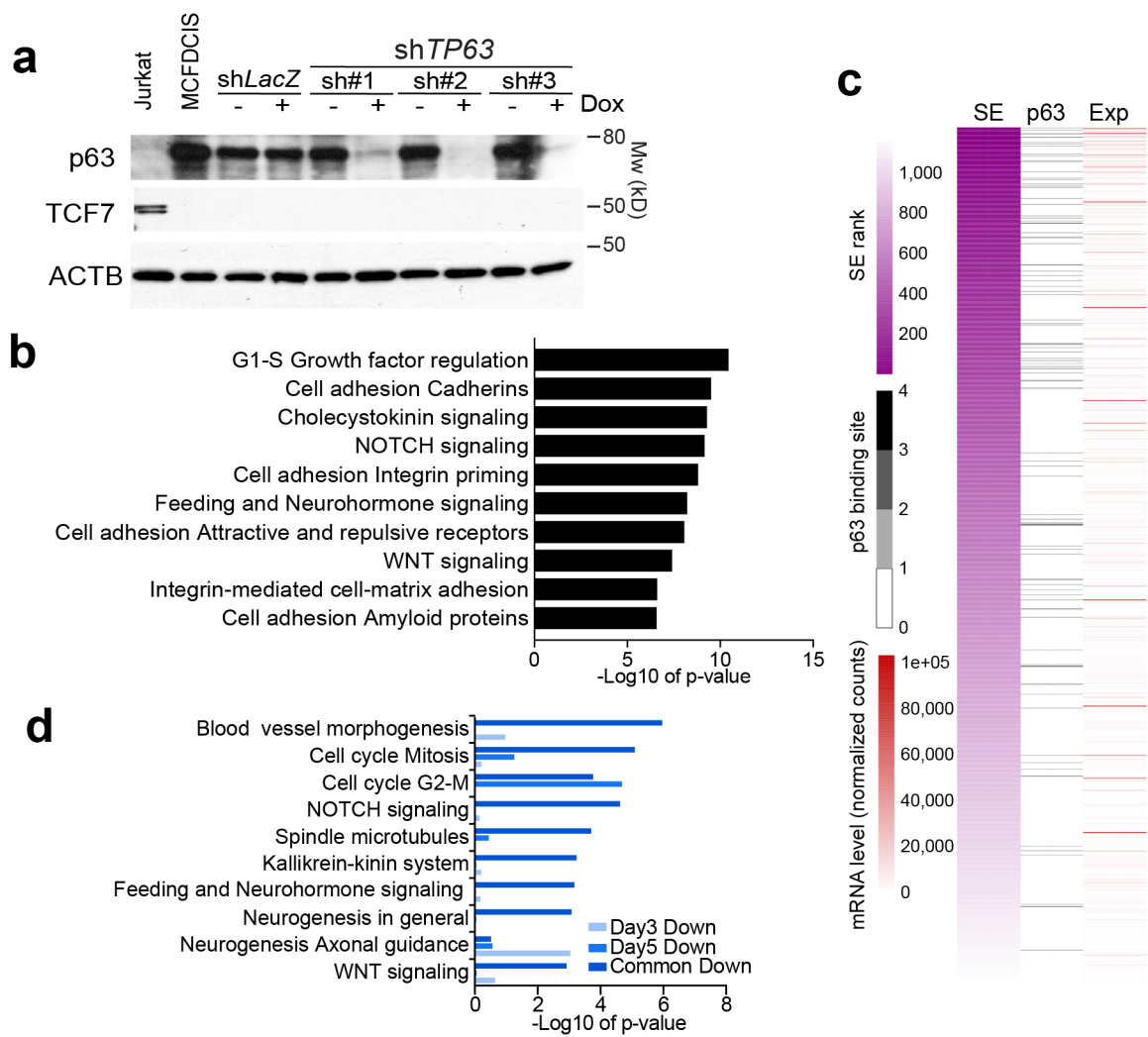
SUPPLEMENTARY FIGURES

Supplementary Figure 1. Heterogeneity of CD10⁺ cell populations in normal breast tissues. **a**, Schematic outline of the purification procedure and phenotype of the cells in cell culture. **b**, Combined analysis of cells for the expression of CD10 and CD44 cell surface markers and ALDH enzymatic activity. Left panel indicates staining of cells for CD10 and CD44 whereas right panel shows ALDH activity of cells selected based on difference in CD10 and CD44 expression. ALDH activity was assessed with and without DAEB (ALDH inhibitor). **c**, Gating strategy used for the quantification of CD10⁺CD44⁺ and CD10⁺CD44⁻ cells. Single cell suspension derived from organoids were stained with viability dye-pacific blue, CD24-APC, CD44-PE, and CD10-FITC. Each single-color control was run in parallel to ensure proper compensation. Cells were gated using forward scatter (FSC) and side scatter (SSC) and then gated for the pacific blue negative viable cells. Next, CD24-APC⁺ cells were gated and sorted for other studies. The remaining CD24-APC negative cells were gated for CD10-FITC⁺CD44-PE⁺ and CD10-FITC⁺CD44-PE⁻ cells and sorted. **d**, viSNE maps of CyTOF analysis of normal breast tissues (CD45 negative population) colored for the indicated markers. Color scale indicates minimum and maximum values of expression in both panels (**d** and **e**). **e**, viSNE maps of CyTOF analysis of myoepithelium (CK8/18 negative, E-Cadherin⁺, SMA⁺ population) colored for the indicated markers. **f**, Heatmap depicting unsupervised clustering of samples based on the expression of top 1,000 most variable genes. Scale indicates relative log₂ fold changes of gene expression, white is median of all samples and red/blue are relative to median of all samples. **g**, Pathways enriched in genes differentially expressed between the indicated groups of samples. Color scale corresponds to -log (p-value) of significance of enrichment, calculated by MetaCore Enrichment Analysis test. **h**, Gene tracks (189.501-189.533Mb) depicting *TP63* isoforms in normal breast and in MCF10DCIS cells. **i**, Gene tracks (133.448-133.486Mb) depicting *TCF7* isoforms in normal breast and in MCF10DCIS cells.

Supplementary Figure 2. Gene expression and p63 and TCF7 binding patterns in DCIS and *BRCA* mutation carriers. **a**, Heatmap depicting unsupervised clustering of normal and DCIS-associated CD10⁺ cells based on the expression of top 1,000 most variable genes. Scale indicates relative log2 fold changes of gene expression, white is median of all samples and red/blue are relative to median of all samples. **b**, Venn diagrams indicating the numbers of p63 and TCF7 binding sites which are unique in control, and in *BRCA1* or *BRCA2* mutation carriers, or overlap between groups.



Supplementary Figure 3. Genomic targets of p63 in MCF10DCIS cells. **a**, Immunoblot analysis of p63 and TCF7 expression in MCF10DCIS cells and three different sh*TP63*-expressing derivatives. Jurkat cells were used as positive control for TCF7. ACTB is loading control. **b**, Pathways enriched in genes associated with p63 and H3K27ac overlapping peaks. **c**, Associations between SE ranking, p63 binding sites, and gene expression levels. **d**, Pathway enrichment analysis of p63 genomic targets downregulated in MCF10DCIS cells after induction of sh*TP63* by Dox.



Supplementary Figure 4. Genomic targets of TCF7 in MCF10DCIS cells. **a**, Gating strategy for the identification of cell cycle distribution. Cells were first gated based on 7-AAD-height/7-AAD-area to identify single cells (gate: singlets). Following, single cells were shown for BrdU-FITC and 7-AAD-A and gated as S-Phase (BrdU-FITC-positive), G1-phase (BrdU-FITC-negative; single 7-AAD content) and G2/M-phase (BrdU-FITC-negative; double 7-AAD content). **b**, Relative fraction of cells in different phases of the cell cycle before and after 5 days of doxycycline treatment to induce shTP63 and TCF7, respectively. **c**, Predicted protein interaction network of TFs identified as core transcriptional regulatory circuits in MCF10DCIS cells expressing TCF7. Legend indicates the source of data used to determine interactions. TFs that are not part of the network were removed. **d**, Heatmap depicting p63 and TCF7 unique peaks and overlapping peaks in MCF10DCIS cells. **e**, Venn diagram indicating the numbers of p63 and TCF7 unique and overlapping binding sites in MCF10DCIS cells. **f**, Dot plot depicting results of cell migration and invasion experiments at different time points following the induction of shTP63 or TCF7 by doxycycline (DOX). **g**, Dot plot depicting results of cell adhesion to different matrices experiments at different time points following the induction of shTP63 or TCF7 by doxycycline (DOX). P-values indicate statistical significance of difference to no DOX determined by t-test. **h**, Simplified model of pathway interactions. Source data are provided as a Source Data file.

