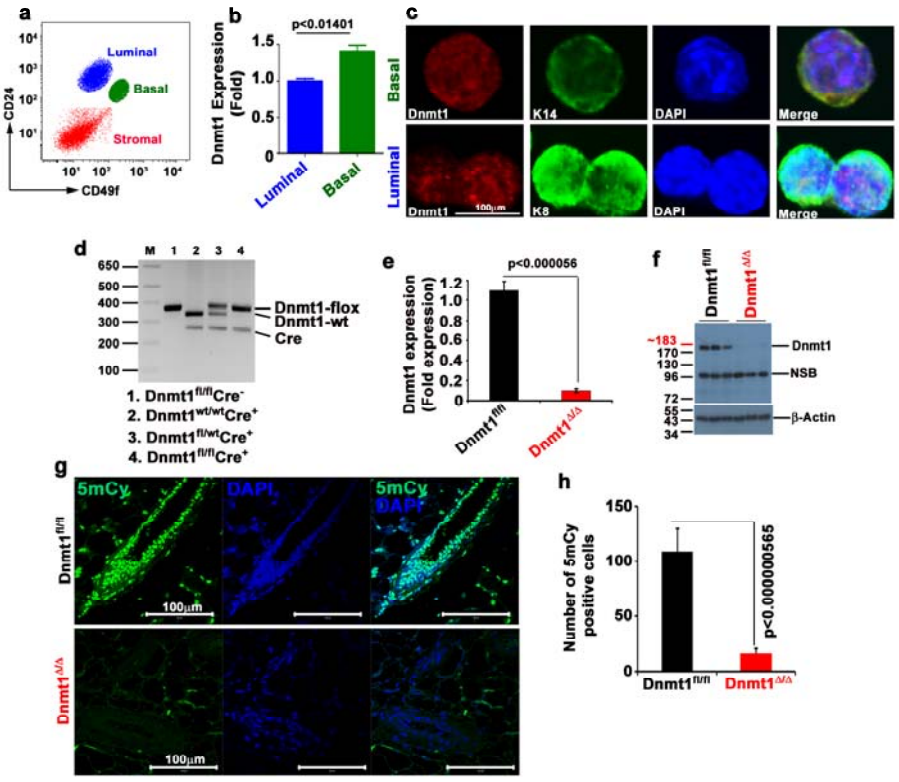
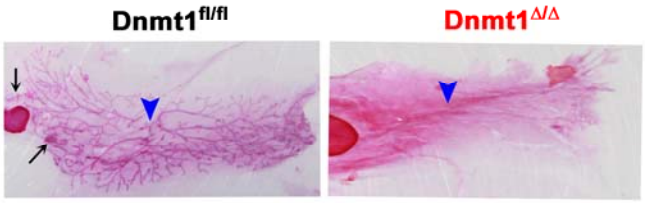
Supplementary Figures



Supplementary Figure 1

Supplementary Figure 1. Confirmation of Dnmt1 conditional knockout out mice. a, Representative images of sorted stem ($Lin^-CD49f^{high}CD24^+$), luminal ($Lin^-CD49f^{low}CD24^+$) and stromal cells ($Lin^-CD49f^-CD24^-$) from normal control mice (n=3). **b,** Real-time PCR analysis data showing the relative levels of Dnmt1 gene transcripts in stem and luminal cells. Data represents Mean ± SD (n= 5 mice). **c,** Representative confocal images of sorted stem cells from normal mammary glands show Keratin 14 (K14, green), Dnmt1 (red) and DAPI (blue) staining. Similarly, the progenitor cells show Keratin 8 (K8, green), Dnmt1 (red) and DAPI (blue) staining. **d,** PCR genotyping analysis of Dnmt1 conditional knockout mice (Lane 4). $Dnmt1^{fl/fl}$-Cre^- mice (lanes 1) were bred with MMTV-Cre (lanes 2) and the resulting $Dnmt1^{fl/wt}$-Cre^+ mice (lane 3) were intercrossed to generate $Dnmt1^{fl/fl}$-Cre^- ($Dnmt1^{fl/fl}$), $Dnmt1^{fl/\Delta}$-Cre+ ($Dnmt1^{fl/\Delta}$) and $Dnmt1^{fl/fl}$-Cre^+ ($Dnmt1^{\Delta/\Delta}$, lane 4) mice. **e,** Real-time PCR analysis shows reduced Dnmt1 transcript in mammary gland of DNMT1-knockout mice compared wild-type mice (n=3). **f**, Representative Western blot analysis shows that ~95% reduced Dnmt1 protein expression in Dnmt1-knockout mice when compared to wild-type. For Western blot analysis, we used mid pregnant (10^{th} day) mice of both wild-type and Dnmt1-knockout mice (n=3). NSB, non-specific band. **g**, Representative confocal images of 5mC (Green), and DAPI (blue) expression in wild-type and $Dnmt1^{\Delta/\Delta}$ mouse mammary gland sections. Scale bar 100 μm. **h,** Number of 5mC positive cells in wild type and $Dnmt1^{\Delta/\Delta}$ mouse mammary gland. Data are shown Mean ± SEM (n=3 mice). Statistical analysis was performed using unpaired Student's *t*-tests.

a



b

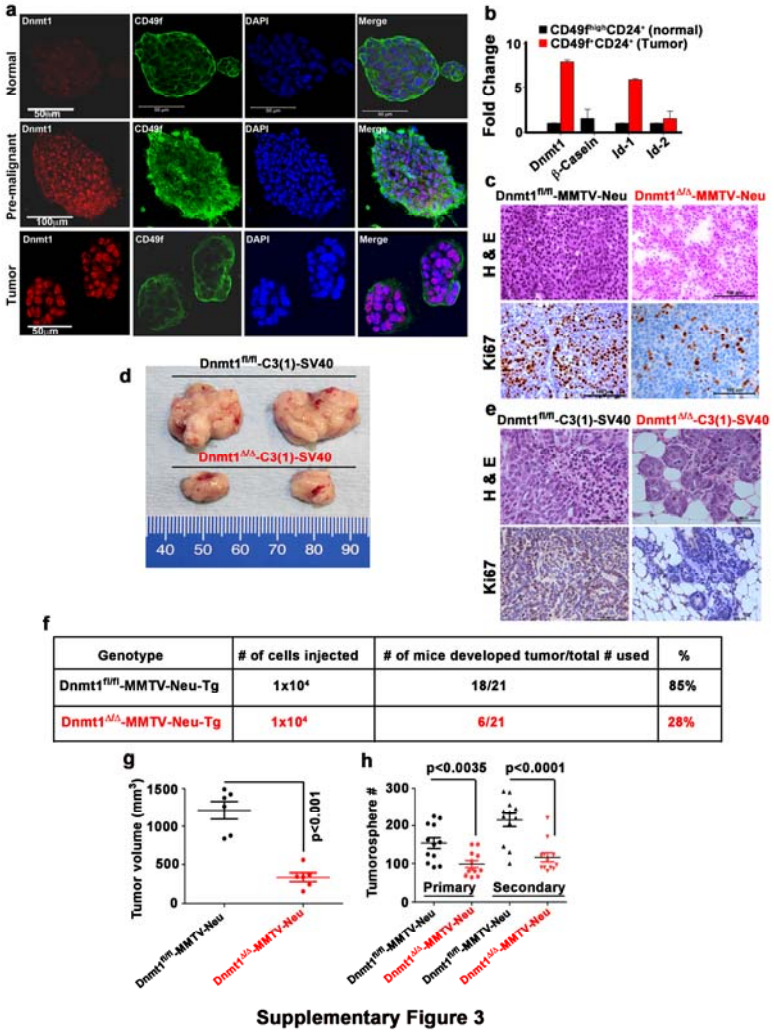
Limiting dilution assay for mammary stem cells isolated from Dnmt1$^{fl/fl}$ and Dnmt1$^{\Delta/\Delta}$ mice

No. of transplanted cells	Dnmt1$^{fl/fl}$	Dnmt1$^{\Delta/\Delta}$
1x10^2	1/4	0/4
5x10^2	3/4	0/4
1x10^3	4/4	0/4
5x10^3	3/3	0/6
1x10^4	3/3	0/4
5x10^4	2/2	2/7
1x10^5	4/4	4/6

Mammary repopulating frequency in total mammary epithelial cell determined by limiting dilution assay was of 1/310 (1/720 - 1/134) from Dnmt1$^{fl/fl}$ and 1/124426 (1/276534 - 1/55986) for Dnmt1$^{\Delta/\Delta}$ mice. CI 95%, p value 8.7e-22

Supplementary Figure 2

Supplementary Figure 2. Dnmt1-deletion abolishes the mammary repopulating unit (MRU) efficiency. a, Representative whole mount mammary gland outgrowth generated from the transplantation of mammary epithelial cells isolated from Dnmt1$^{fl/fl}$ and Dnmt1$^{\Delta/\Delta}$ mice into the cleared fat pads. Blue arrows indicate site of injected cells and black arrows indicate cisternal duct. **b,** Table shows the frequency of mammary repopulating unit (MRU) in total mammary epithelial cell population as determined by limiting dilution assay.

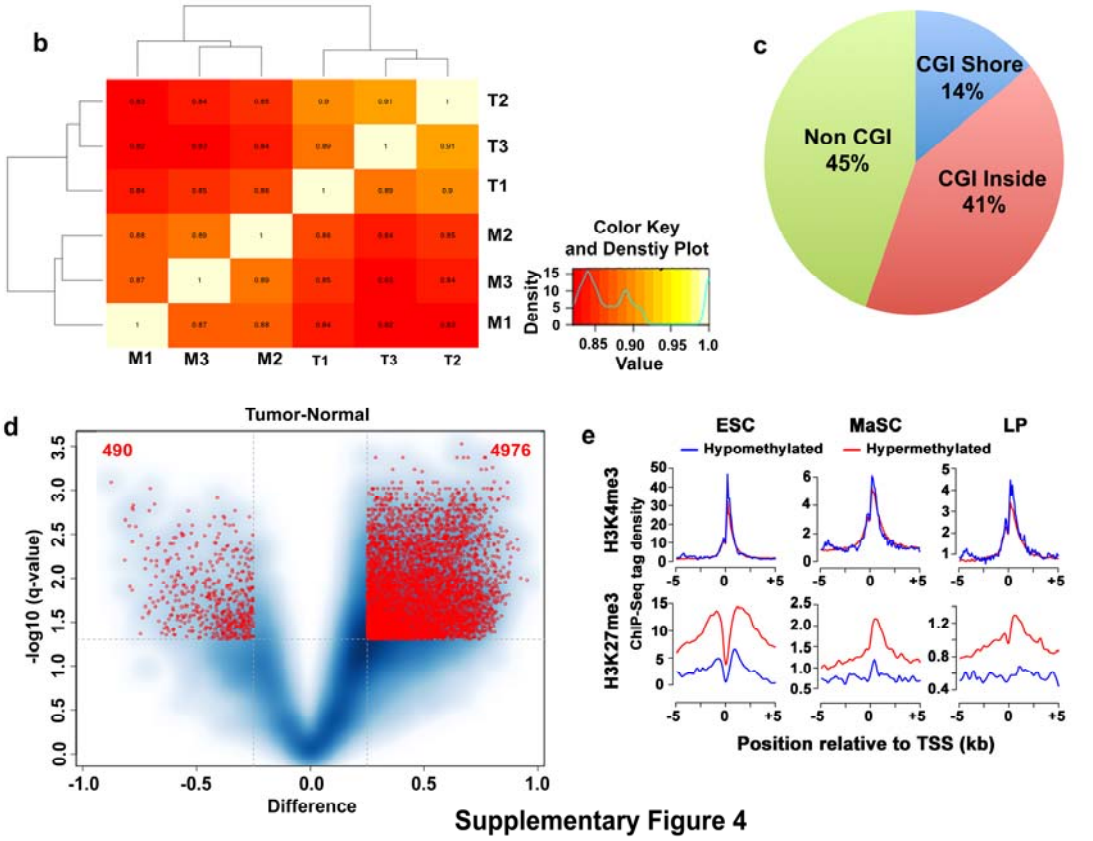


Supplementary Figure 3

Supplementary Figure 3. DNMT1 deletion suppresses mammary tumorigenesis by limiting CSCs a, Representative mammosphere from normal control mice and tumorospheres from premalignant and tumor tissues of MMTV-Neu-Tg mice showing Dnmt1 (red), CD49f (green) and DAPI (blue) expression. **b,** Real-time PCR analysis for the Dnmt1 in normal stem cells (n=5 mice) and tumor-propagating cells (n=3 mice). **c,** Representative H & E and Ki67 staining of tumor tissues of Dnmt$^{fl/fl}$-MMTV-Neu-Tg and Dnmt1$^{\Delta/\Delta}$-MMTV-Neu-Tg mice. Images were taken at 40x magnification. **d,** Representative tumor images of Dnmt$^{fl/fl}$-C3(1)-SV40-Tg and Dnmt1$^{\Delta/\Delta}$-C3(1)-SV40-Tg mice. **e,** Representative H & E and Ki67 staining of tumor tissues of Dnmt$^{fl/fl}$-C3(1)-SV40-Tg and Dnmt1$^{\Delta/\Delta}$-C3(1)-SV40-Tg mice. Images were taken at 40x magnification. **f**, Number of tumor developed by injecting 10,000 cells, which derived from the primary tumorospheres generated from Dnmt$^{fl/fl}$-MMTV-Neu-Tg and Dnmt1$^{\Delta/\Delta}$-MMTV-Neu-Tg mice (8 month old premalignant mice with no gross visible tumors), into mammary gland of NOD/SCID mice. **g,** Tumor volume and **h**, Primary and secondary tumorospheres from premalignant Dnmt$^{fl/fl}$-MMTV-Neu-Tg and Dnmt1$^{\Delta/\Delta}$-MMTV-Neu-Tg mice. Statistical analysis was performed using unpaired Student's *t*-tests. Figures represent Mean ± SEM.

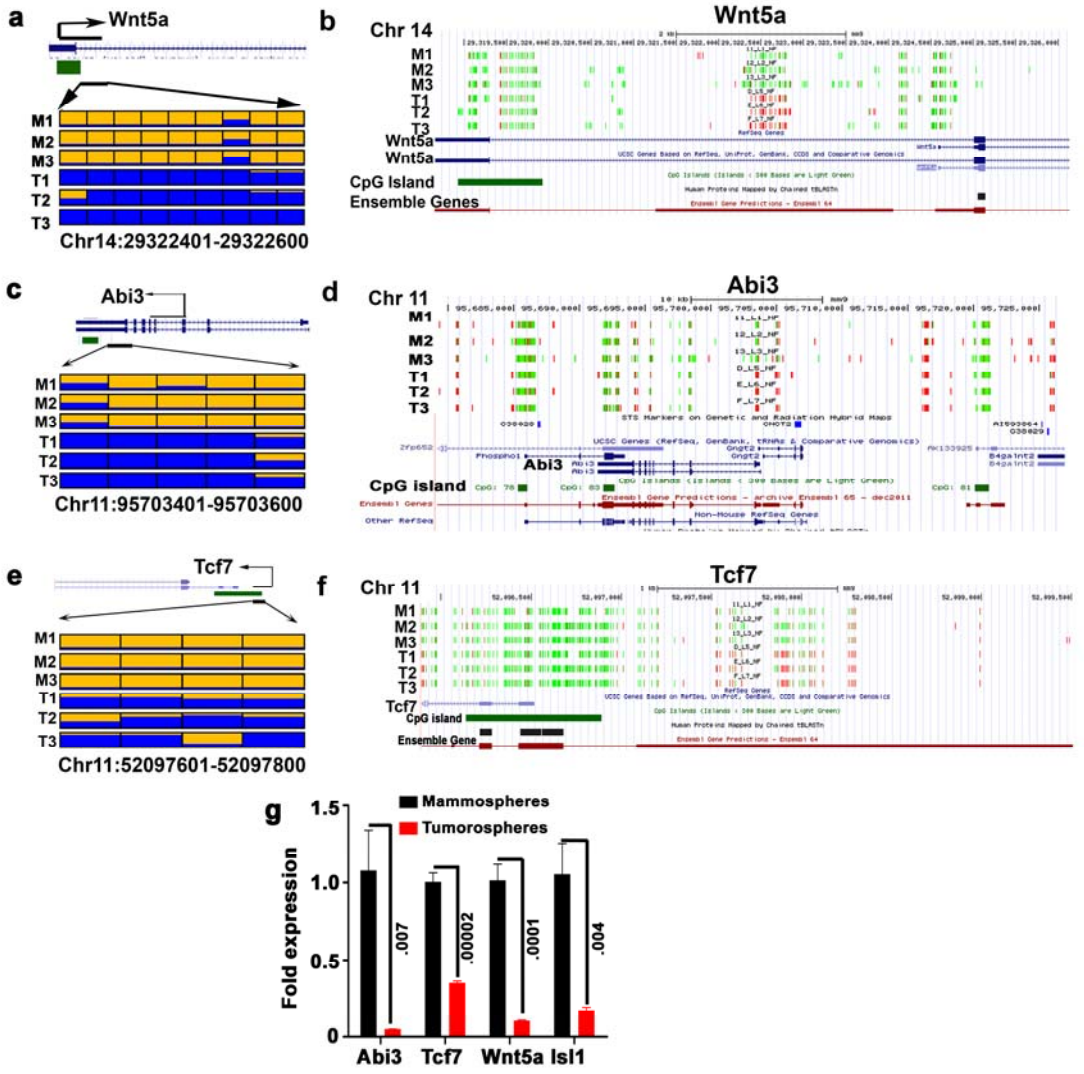
a

Sample	# of Reed	Mapping rate(%)	# of CpGs(>0 read)	#of CpGs(>5 read)	Av. CpG methylation	Av. CpG coverage
M1	28338880	84.81	1814591	1122587	0.420016388	65.0125
M2	33042723	90.28	2285366	1324628	0.376114535	77.2136
M3	30716473	78.21	3017338	1284662	0.321745066	81.6481
T1	29385183	89.51	2018959	1226601	0.455443737	69.8052
T2	34408486	93.59	1906201	1371104	0.397755796	99.7525
T3	35800013	90.51	1902059	1294674	0.392724546	107.1733



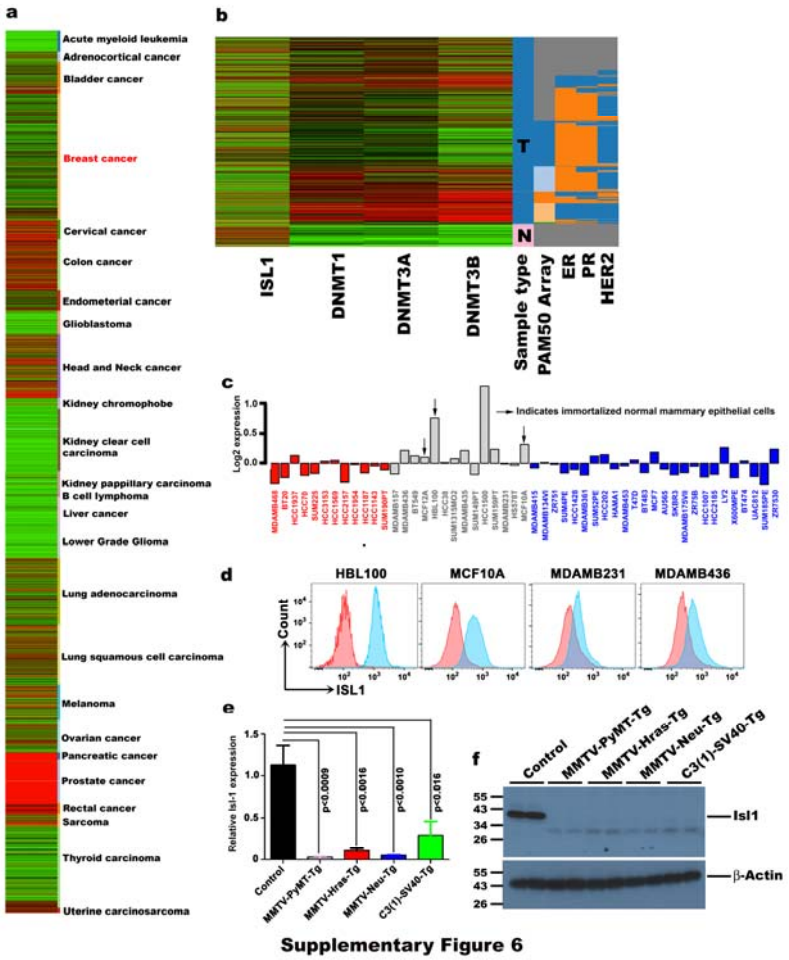
Supplementary Figure 4

Supplementary Figure 4. Genome-scale DNA methylation studies in mammospheres and tumorospheres. a, Summary statistics of DNA methylation. **b,** Correlation coefficients matrix comparing DNA methylation between mammospheres and tumorospheres. A high similarity was observed among mammospheres M1, M2 and M3 and tumorospheres T1, T2 and T3 samples. **c,** Genome-wide distribution of 2341 DMRs that were identified in tumorospheres. **d,** Plot of methylation difference (x-axis) vs. significance (y-axis, FDR q-value) shows the number of DMRs in tumorospheres compared to mammospheres. Red circles mark DMRs that reach both criteria for differential methylation in our analysis (methylation difference >0.25 and FDR q-value < 0.01. **e,** ChIP-Seq read density plots showing the levels of H3K4me3 and H3K27me3 in ESCs, MaSCs, and luminal progenitors (LPs) at promoters of genes hypermethylated (red) and hypomethylated (blue) in tumorspheres. TSS, transcription start site.



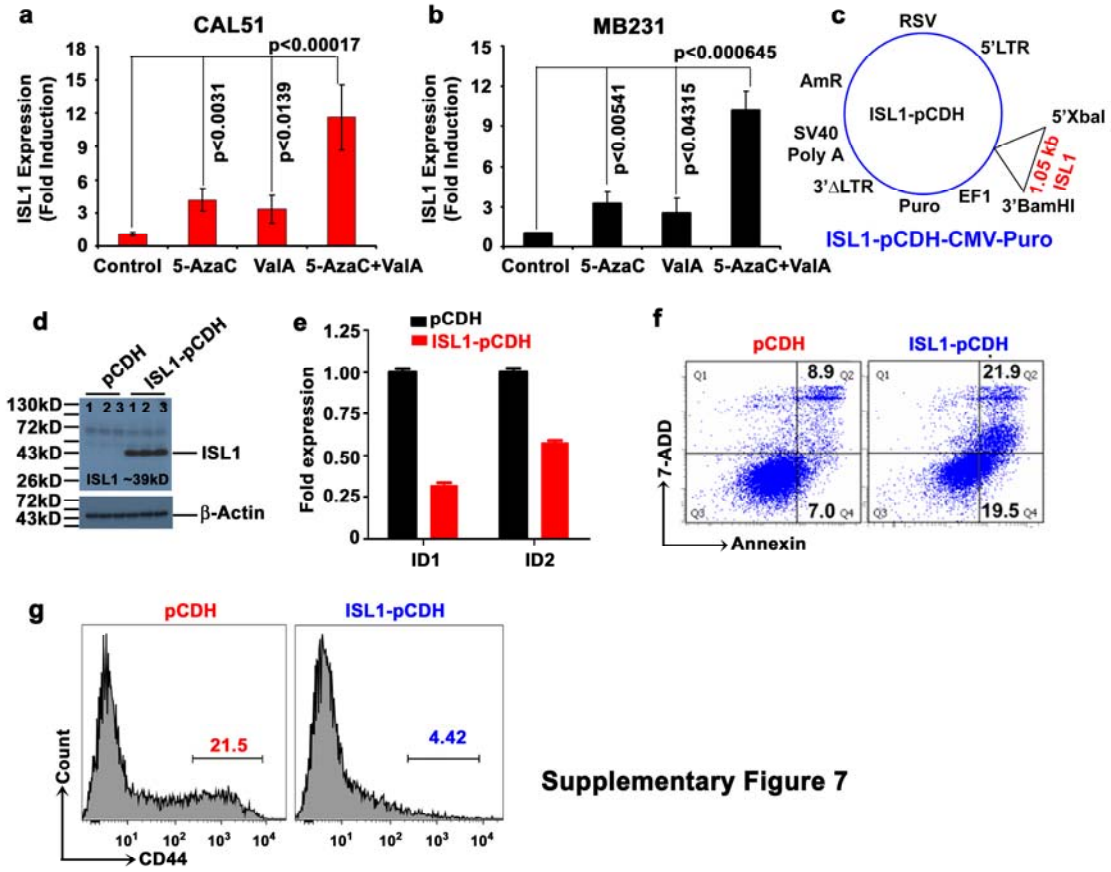
Supplementary Figure 5

Supplementary Figure 5. Wnt5a, Abi3, and Tcf7 are hypermethylated and downregulated in tumorospheres. a, c, & e, Single-base CpG methylation patterns in mammospheres and tumorospheres for Wnt5a, Abi3 and TCF7 respectively. Each box represents a CpG site. The color indicates the level of methylation, blue indicates hypermethylation and yellow indicates unmethylation. **b, d, f,** UCSC genome browser screenshot showing the RRBS results for Wnt5a, Abi3, and Tcf7. Red and green colors indicate methylated and unmethylated CpG sites, respectively. **g**, Real time PCR relative expression of Abi3, Tcf7, Wnt5a, and Isl1 expression in Mammospheres and Tumorospheres generated from normal mammary gland and MMTV-Neu tumors (n=4). Statistical analysis was performed using unpaired Student's *t*-tests. Bar diagram represents Mean ± SD.



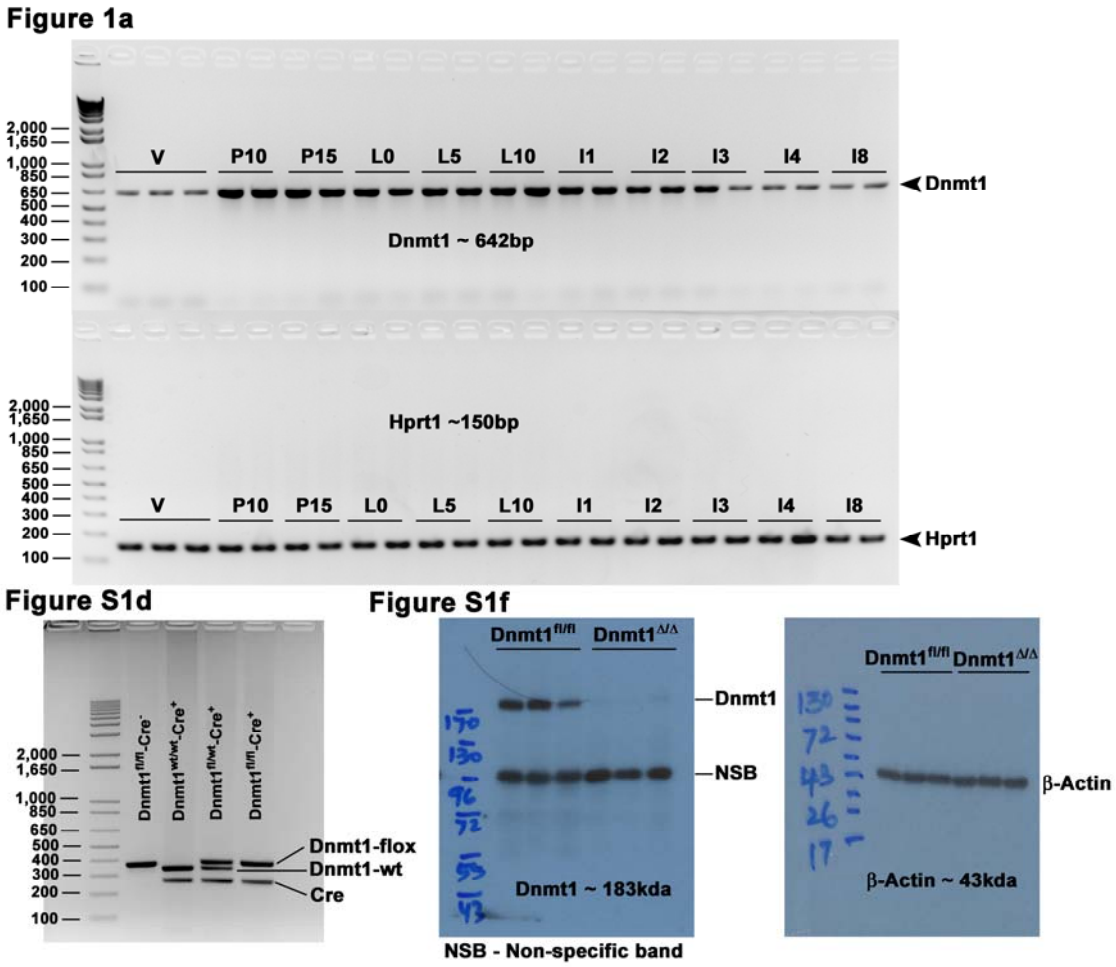
Supplementary Figure 6

Supplementary Figure 6. ISL1 expression is downregulated in human primary tumor tissues and in breast cancer cell lines . a, The Pan Cancer UCSC Cancer Genomics heatmap shows gene expression for ISL1 in different types of cancer, over a set of genes differentially expressed. Red color indicating up-regulation and green color indicating down-regulation. **b,** Cancer UCSC Cancer Genomics heatmap shows gene expressions of ISL1, DNMT1, DNMT3A and DNMT3B across the 1160 patients' database in TCGA breast invasive carcinoma (RNA Seq). Red color indicating up-regulation and green color indicating down-regulation. **c,** Gene Expression of ISL1 across the 51 individual breast cancer cell lines searched in GOBO online tool. Color codes represents basal A (red), basal B (grey) and luminal (blue) type breast cancer subgroups. Arrow indicated immortalized normal mammary epithelial cell line. **d,** Expression of ISL1 protein in HBL100, MCF10A, MDAMB231 and MDAMB436 cell lines. Red color = Unstained cells (Isotype IgG, PE Control), Blue color= ISL1 positive cells. **e,** Expression of Isl1 mRNA in normal mouse mammary gland (n =3), MMTV-PyMT (n =3), and MMTV-Hras (n =3), MMTV-Neu (n =3), and C3(1)-SV40 analyzed by q-PCR. **f,** Isl1protein expression in normal control mouse mammary gland (n =2), MMTV-PyMT (n =2), MMTV-Hras (n =2), MMTV-Neu (n =2), and C3(1)-SV40 (n=2) analyzed by western blot using Isl1 antibody. β-actin was used as a loading control. Statistical analysis was performed using unpaired Student's *t*-tests. Bar diagram represents Mean ± SEM.



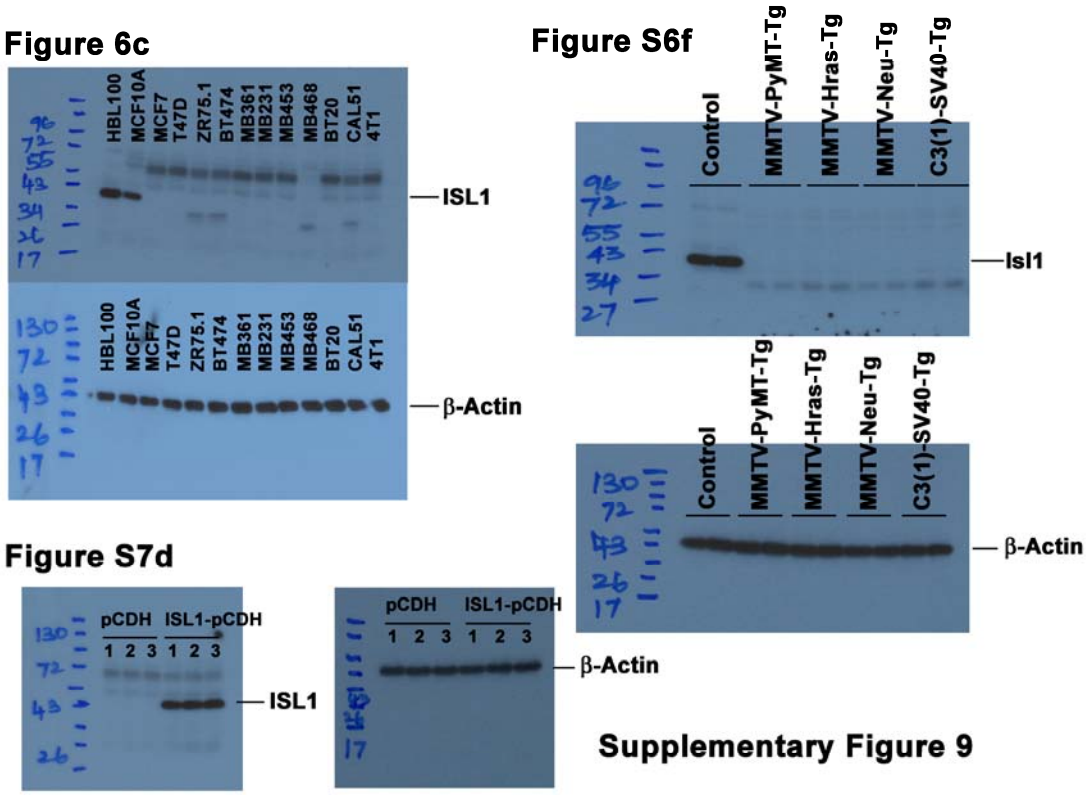
Supplementary Figure 7

Supplementary Figure 7. Inhibition of DNMT activity or ISL1 expression in cancer cells reduces colony formation, induces apoptosis, blocks cell migration and inhibits cancer stem cell marker expression. a-b, DNA methylation (DNMT) and histone acetyl transferase (HDACs) inhibitor reactivates ISL1 expression in human breast cancer cells. Human breast cancer cell lines (CAL51 and MB231) were treated with DNMT inhibitor 5-azacytidine (5-AzaC, 1 μg/ml) and HDAC inhibitor valproic acid (ValA, 0.1 mg/ml) for 72 h. Cells were collected at the end of the treatment period and RNA was extracted using TRIZOL reagent. ISL1 mRNA expression was analyzed by quantitative RT-PCR (qPCR) using human ISL1 specific primers. HPRT was used as an internal control. **c,** ISL1-pCDH vector map. ISL1 ORF (1.05kb) was subcloned into ISL1-pCDH-CMV-Puro vector in BamHI/Xbal cloning site. **d,** Western analysis of protein samples, derived from CAL51-pCDH and CAL51-ISL1-pCDH stable cell lines, shows ISL1 expression. n=3 independent experiments. β-actin was used as a loading control. **e,** Real time PCR relative expression of ID1 and ID2 genes in CAL51-pCDH and CAL51-ISL1 expressing stable cells. **f**, Stable expression of ISL1 in human breast cancer cell line induces apoptosis assessed by FACS analysis of FITC Annexin staining. **g,** ISL1 stable expression in human breast cancer cell reduced the expression of the cancer stem cells marker CD44 as analyzed by FACS. Data represents Mean ± SD from three independent experiments.



Supplementary Figure 8

Supplementary Figure 8. Original images of gel pictures. Figure 1a, a representative gel image of DNA methyltransferase 1 (Dnmt1) gene transcript in various stages of mammary gland development using mouse specific Dnmt1 primers (top). Hypoxanthine phosphoribosyltransferase 1 (Hprt1) was used as a loading control (bottom) (n=2-3 mice per time point). **Figure S1d,** a representative gel image of Dnmt1 genotype using primers specific for wild-type ($Dnmt1^{fl/fl}$-Cre^{-}) and Dnmt1-knockout ($Dnmt1^{fl/fl}$-Cre^{+}) mice. **Figure S1f,** representative western blot image of Dnmt1 protein expression in wild-type ($Dnmt1^{fl/fl}$) and Dnmt1-knockout ($Dnmt1^{\Delta/\Delta}$) mice using anti-Dnmt1 antibody (left). β-actin was used as a loading control (right) (n=3).



Supplementary Figure 9. Original images of gel pictures. Figure S6c, representative western blot images of ISL1 (top) and β-actin (bottom) in the respective cell lines using anti-ISL1 and anti-β-actin antibodies, respectively. **Figure S6f,** representative western blot images of Isl1 (top) and β-actin (bottom) in normal mammary tissue (control) and mammary tumor tissues of respective transgenic mice (MMTV-PyMT, MMTV-Hras-Tg, MMTV-Neu-Tg, and C3(1)-SV40-Tg) using anti-Isl1 and anti-β-actin antibodies, respectively (n=2). **Figure S7d,** representative western blot images ISL1 (left) and β-actin (right) in CAL51-pCDH and CAL51-ISL1 stable cell lines using anti-ISL1 and anti-β-actin antibodies, respectively (n=3).

Supplementary Methods

List of primers used for the semi-quantitative analysis:

1. Mouse Dnmt1: 5’-AGTGTGGCCAGCACCTAGAC-3’(forward)
   5’-CCTTGGCTTCGTCGTAACTC-3’ (reverse)
2. Mouse Hprt1: 5’-GCGTCGTTAGCGATGATGAAC-3’(forward)
   5’-CCTCCCATCTCCTTCATGACATCT-3’ (reverse)

List of primers used for the Real-Time PCR analysis:

1. Mouse Dnmt1:	5’-CCTGGCTAAAGTCAAGTCCCT-3’(forward) 5’-GTGTGTGTTCCGTTCTCCAAG-3’ (reverse)
2. Mouse Id1:	5’-CCTAGCTGTTCGCTGAAGGC-3’(forward) 5’-CTCCGACAGACCAAGTACCAC-3’ (reverse)
3. Mouse Id2:	5’-TCCGGTGAGGTCCGTTAGG-3’(forward) 5’-CAGACTCATCGGGTCGTCC-3’ (reverse)
4. Mouse β-Casein :	5’-TTCAGAAGGTGAATCTCATGG GA-3’(forward) 5’-CTGGATGCTGGAGTGAACTT-3’ (reverse)
5. Mouse Abi3:	5’-CAGGTTACCCACTTGGTAGGC-3’(forward) 5’-CTACTGCGAGGATAACTACTTGC-3’ (reverse)
6. Mouse Wnt5a:	5’-CAACTGGCAGGACTTTCTCAA-3’(forward) 5’-CCTTCTCCAATGTACTGCATGTG-3’(reverse)
7. Mouse TCF7:	5’-TCATCACGTACAGCAATGAACA-3’(forward) 5’-CGACAGCGGGTAATATGGAGAG-3’(reverse)

8. Human ID1:	5’-CCTAGCTGTTCGCTGAAGGC-3’(forward) 5’-CTCCGACAGACCAAGTACCAC-3’ (reverse)
9. Human ID2:	5’-TCCGGTGAGGTCCGTTAGG-3’(forward) 5’-CAGACTCATCGGGTCGTCC-3’ (reverse)
10. Mouse Isl1:	5’-TTTCCCTGTGTGTTGGTTGC-3’(forward) 5’-TGATTACACTCCGCACATTTCA-3’ (reverse)
11. Human ISL1:	5’-GCGGAGTGTAATCAGTATTTGGA-3’(forward) 5’-GCATTTGATCCCGTACAACCT-3’ (reverse)