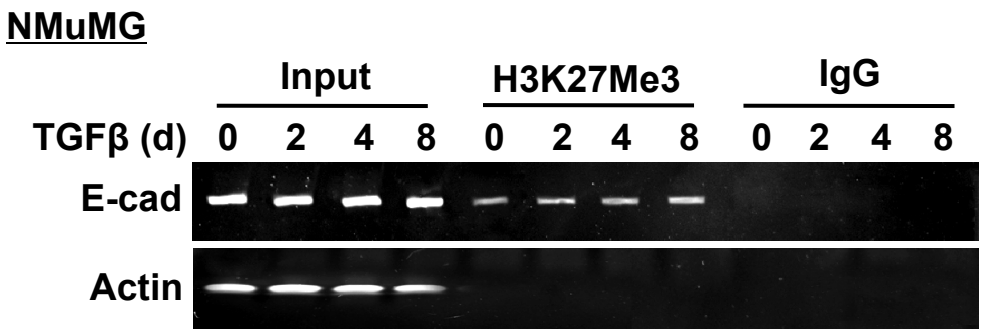
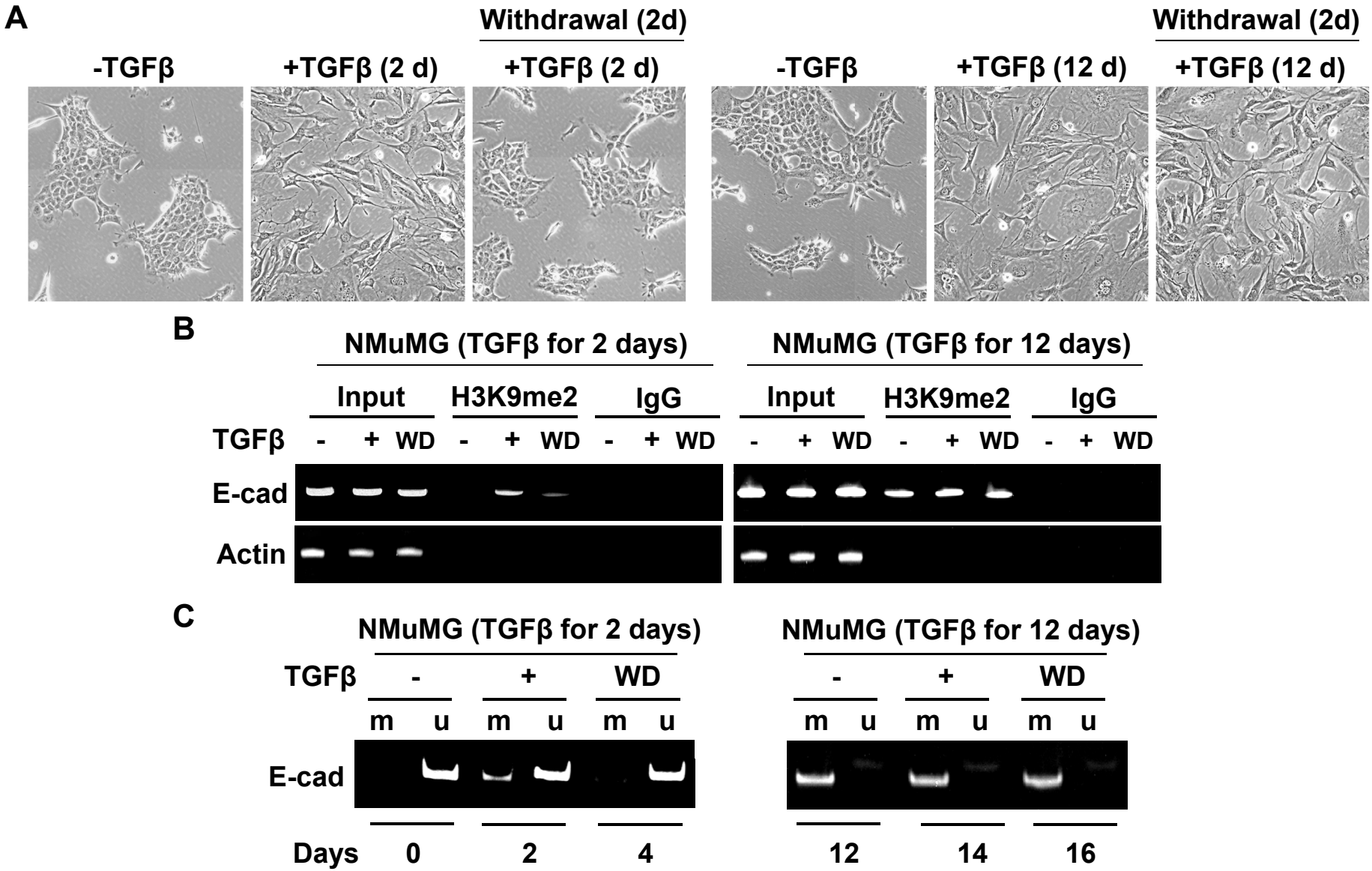
Supplementary Table S1. Tumor samples used for analysis

Sample#	Age	Tumor size (cm)	pN-Stage	Stage	BNG (grade)	ERα	PR	HER2 (FISH)	Triple negative
T1	46	3	N1a	III	2	Pos	Neg	N	
T2	58	1	N0(i-)	I	3	Pos	Neg	N	
T3	51	2	N1mi	II	2	Pos	Pos	N	
T4	74	2	N0	II	3	Pos	Neg	Y	
T5	33	3	N3a	III	2	Pos	Pos	N	
T6	52	3	N3c	III	2	Neg	Neg	Y	
T7	68	2	N1a	II	3	Pos	Pos	N	
T8	47	2	N1mi	II	2	Pos	Neg	Y	
T9	72	3	N0	IV	3	Neg	Neg	Y	
T10	68	2	N1mi	II	2	Pos	Pos	N	
T11	67	2	N1a	II	3	Pos	Pos	N	
T12	53	3	N2a	III	3	Pos	Neg	N	
T13	55	2	N2a	III	2	Pos	Neg	N	
T14	54	2	N0(i-)	II	2	Pos	Pos	N	
T15	52	3	N3a	III	2	Pos	Pos	N	
T16	51	3	N0	II	2	Pos	Neg	N	
T17	32	4	N1	III	3	Neg	Neg	N	yes
T18	67	3	N3a	III	2	Neg	Neg	N	yes
T19	55	4	N1a	III	3	Neg	Neg	N	yes
T20	40	3	N3a	III	2	Neg	Neg	N	yes
T21	62	2	N0	II	2	Neg	Neg	N	yes
T22	46	4	N1a	III	3	Neg	Neg	N	yes
T23	65	2	N0(i+)	II	3	Neg	Neg	N	yes
T24	53	3	N2a	III	3	Neg	Neg	N	yes
T25	50	2	N1mi	II	2	Pos	Pos	N	
T26	63	4	N1a	IV	3	Pos	Pos	Y	
T27	71	2	N0(i-)	II	2	Pos	Pos	N	
T28	30	2	N1a	II	3	Pos	Pos	N	
T29	52	4	N3	III	3	Neg	Neg	N	yes
T30	47	2	N3b	III	3	Neg	Neg	N	yes
T31	51	4	N3a	III	3	Neg	Neg	N	yes
T32	40	3	N3a	III	2	Neg	Neg	N	yes
T33	82	3	N0	II	3	Neg	Neg	N	yes
T34	54	4	N3c	III	3	Neg	Neg	N	yes
T35	47	2	N3a	III	3	Neg	Neg	N	yes
T36	29	2	N0	II	3	Neg	Neg	N	yes
T37	70	2	N0	II	2	Pos	Neg	N	
T38	46	4	N0(i-)	III	3	Neg	Neg	Y	
T39	46	4	N3	III	3	Pos	Pos	N	
T40	64	4	N3b	III	3	Pos	Neg	N	
T41	63	1	N1	II	3	Pos	Pos	N	



Supplementary Figure S1. TGF-β-induced EMT does not significantly alter the level of H3K27me3 at the E-cadherin promoter. NMuMG cells were treated with TGF-β1 (5 ng/ml) for different time intervals, and H3K27me3 at the E-cadherin promoter was analyzed by ChIP.

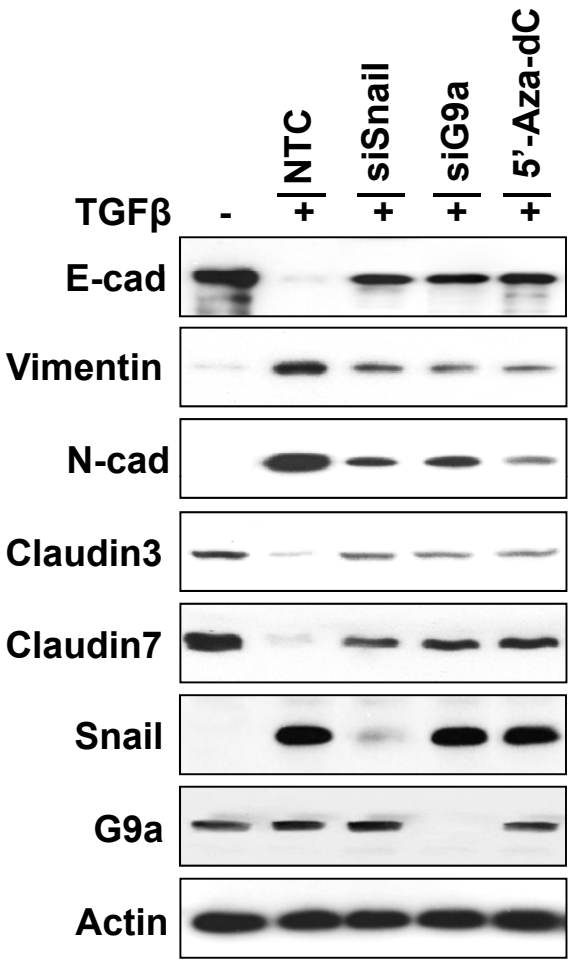


Supplementary Figure S2. Changes of cellular morphology, H3K9me2 and DNA methylation on the E-cadherin after TGF-β withdrawal from NMuMG cells treated with TGF-β for 2 and 12 days, respectively.

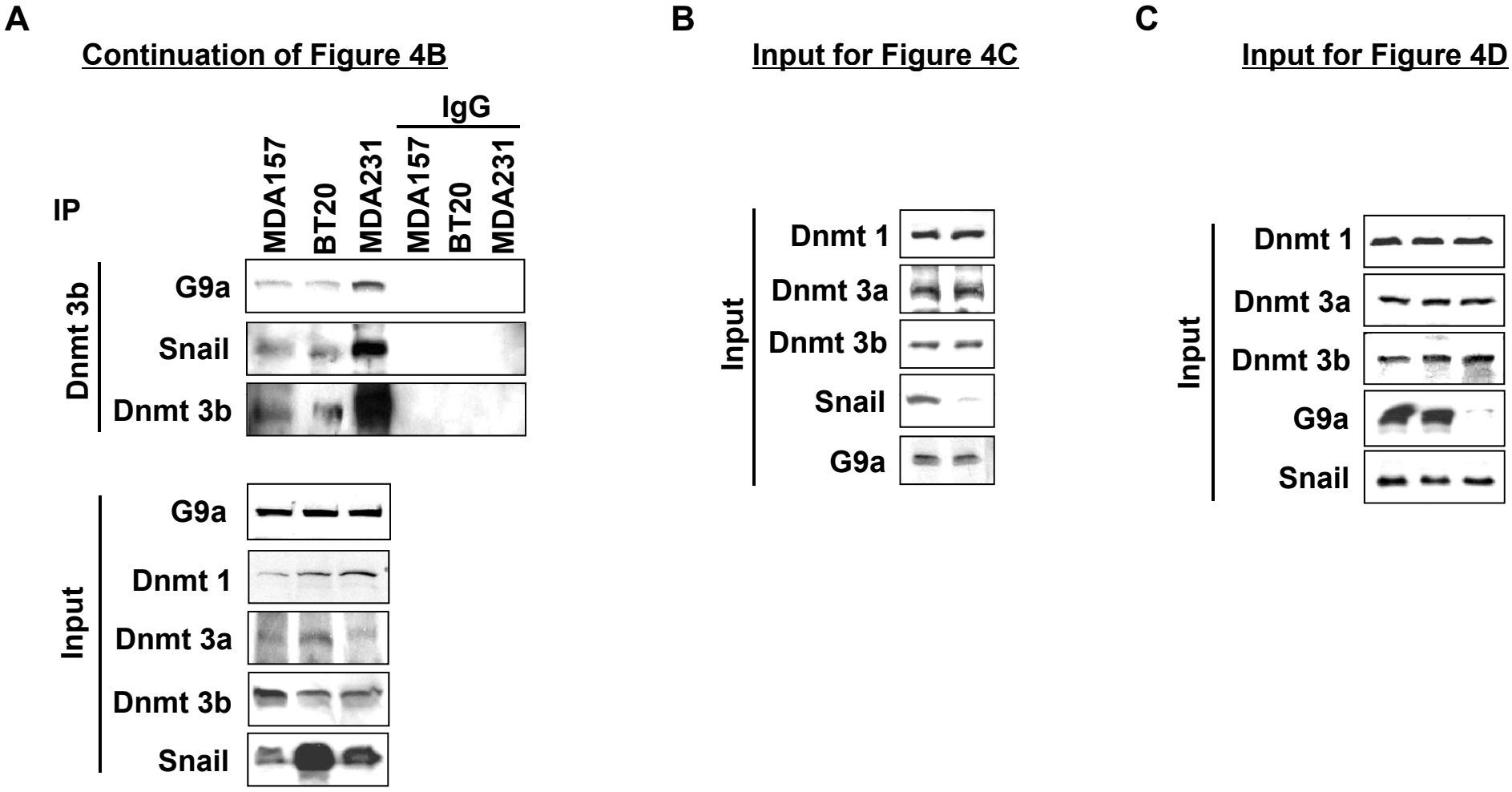
(**A**) NMuMG cells were treated with TGF-β for 2 and 12 days, respectively, followed by TGF-β withdrawal for an additional 2 days. The morphological changes of these cells are shown in the phase contrast images.

(**B**) Cells were treated as described above, H3K9me2 on the E-cadherin promoter was analyzed by ChIP assay.

(**C**) Cells were treated as described above, DNA methylation on the E-cadherin promoter was examined by MSP analysis.



Supplementary Figure S3. G9a, Snail or non-target control (NTC) siRNA was expressed in NMuMG cells followed by TGF-β treatment for 2 days or cells were treated with 5'-Aza-dC. Expression of E-cadherin, Vimentin, N-cadherin, Claudin 3 and 7, as well as Snail and G9a were examined by Western blotting. Actin was used as a loading control.

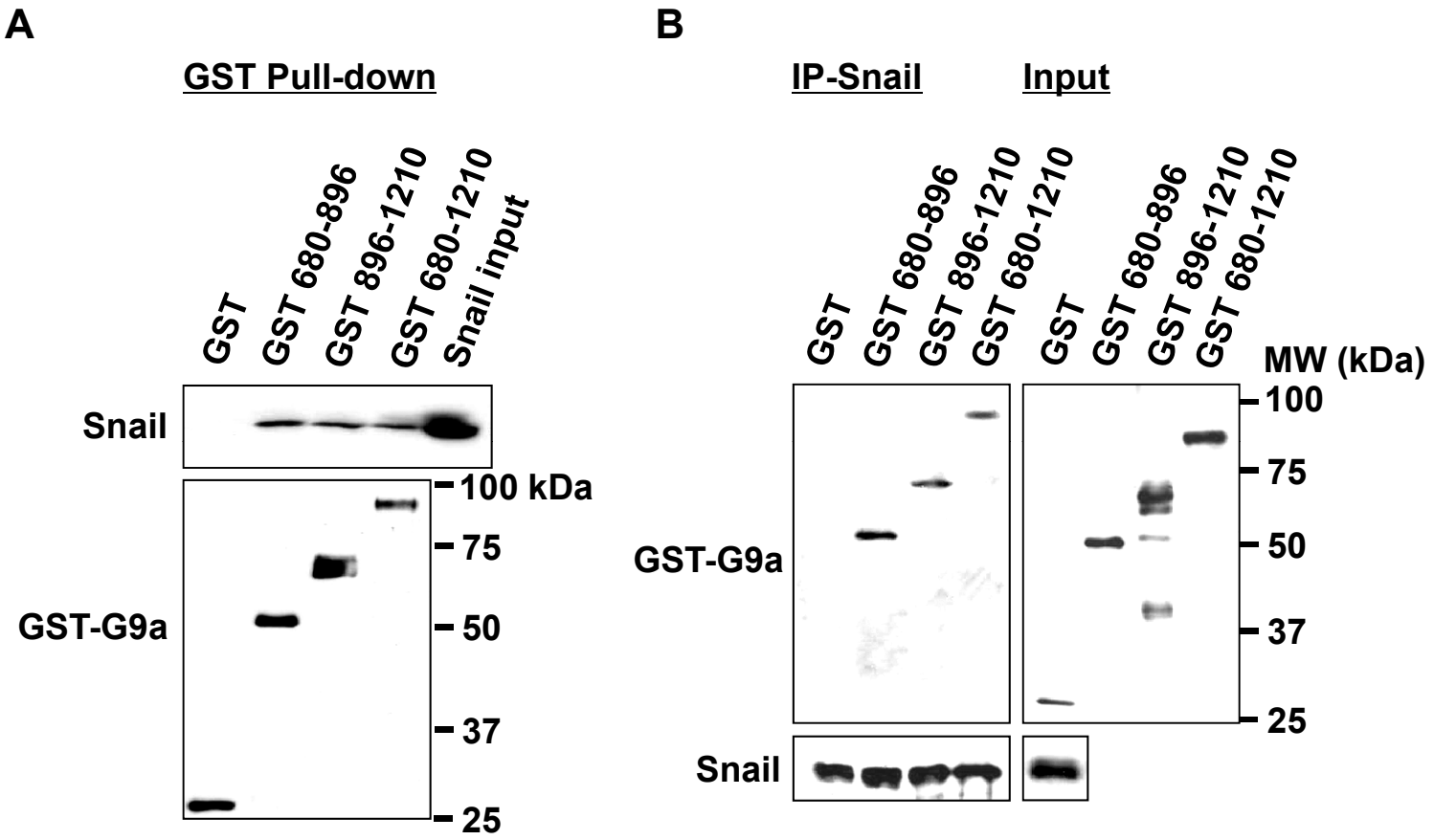


Supplementary Figure S4. Snail forms a complex with G9a and Dnmts.

(**A**) Snail forms a complex with G9a and Dnmts. Endogenous Dnmt3b was immunoprecipitated from MDA-MB157, BT20 and MDA-MB231 cells, and bound endogenous Snail and G9a were examined by Western blotting. Input lysate used for the immunoprecipitation (Fig. 4B) was shown on the bottom panel.

(**B**) Input lysate used for the immunoprecipitation in Figure 4C was shown.

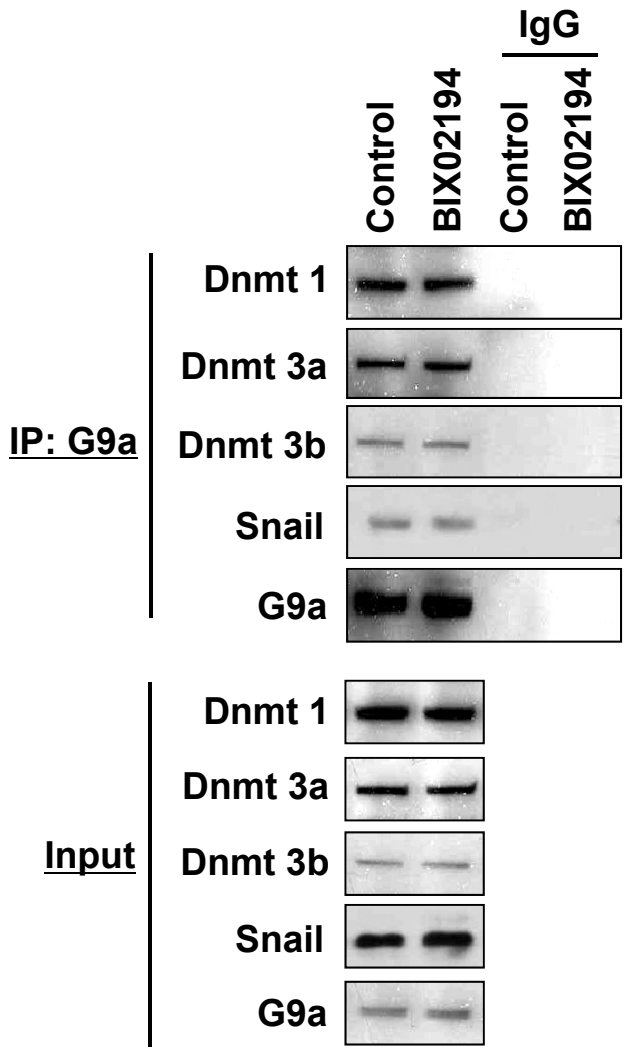
(**C**) Input lysate used for the immunoprecipitation in Figure 4D was shown.



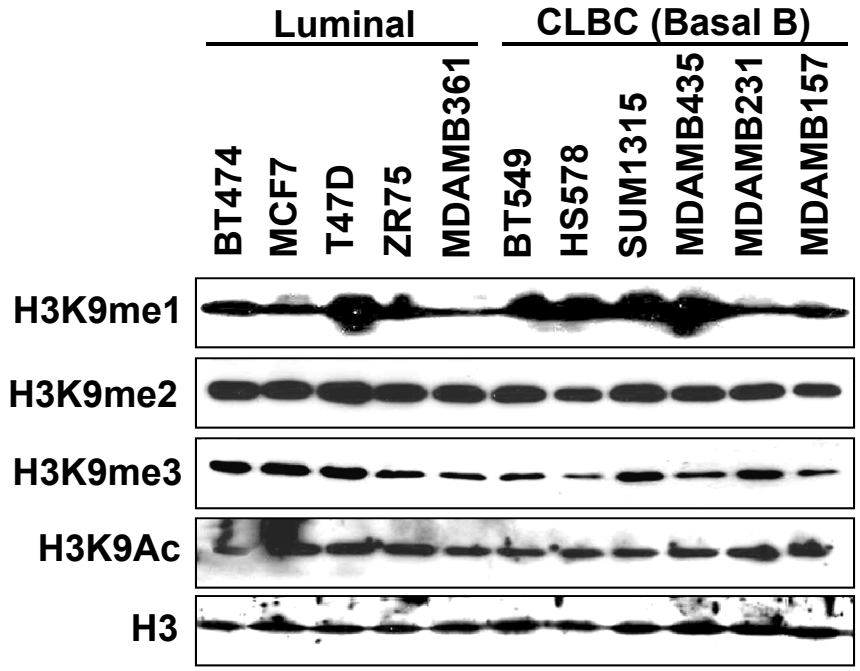
Supplementary Figure S5. Snail interacts directly with G9a.

(**A**) Full-length Snail was purified as a bacterially expressed GST fusion protein followed by excision and removal of the GST protein using TEV protease**.** Purified full-length Snail was incubated with various GST-G9a deletion mutants. After G9a was pulled-down, bound Snail was examined by Western blotting.

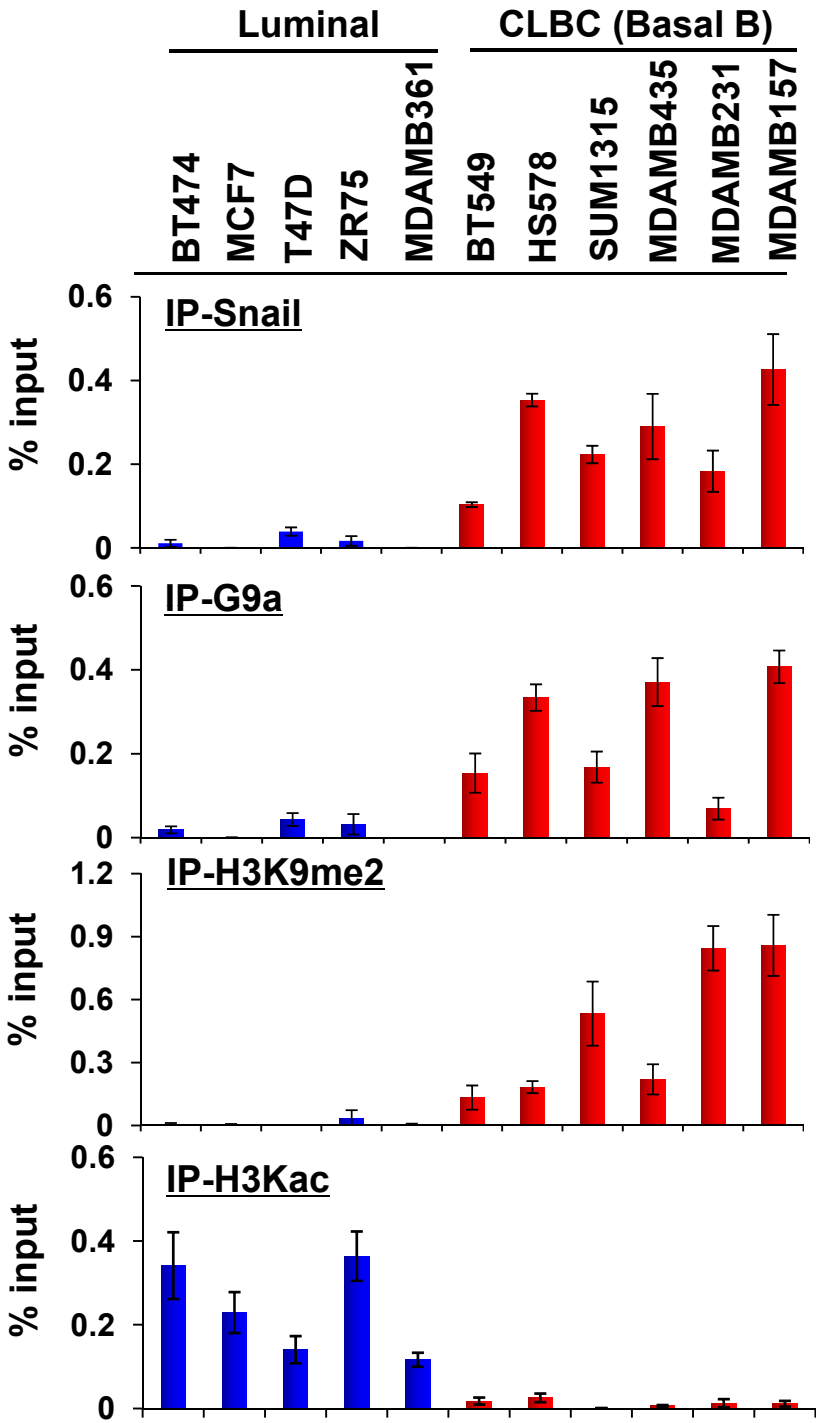
(**B**) Purified full-length Snail was incubated with the various deletions of GST-G9a described above. After Snail was immunoprecipitated, the bound G9a deletion mutants were examined by Western blotting.



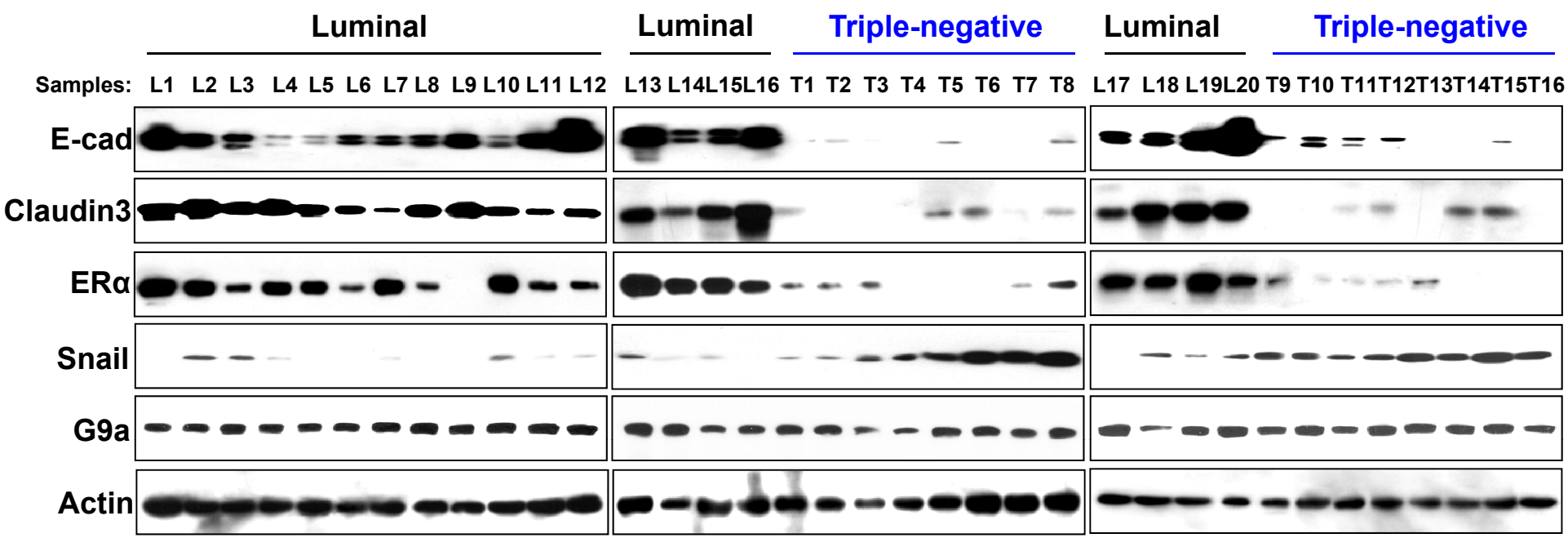
Supplementary Figure S6. Inhibition of G9a activity does not alter the interaction of G9a with Snail and Dnmts.
The G9a inhibitor BIX01294 (2.5 μM) was added to the freshly-prepared MDA-MB157 lysate during immunoprecipitation. After immunoprecipitating endogenous G9a, bound Snail and Dnmts were analyzed by Western blotting. Input lysate used for the immunoprecipitation was shown on the bottom panel.

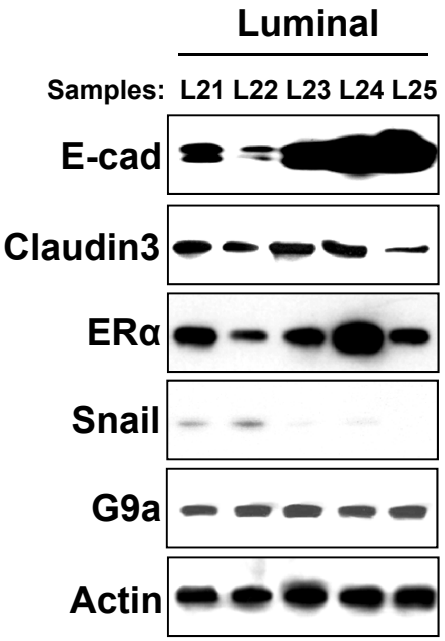


Supplementary Figure S7. There is no significant difference in global H3K9 methylation and acetylation between luminal and CLBC cell lines. Cell extracts were prepared from five luminal and six CLBC cell lines, and the global level of H3K9 methylation and acetylation were analyzed by Western blotting.

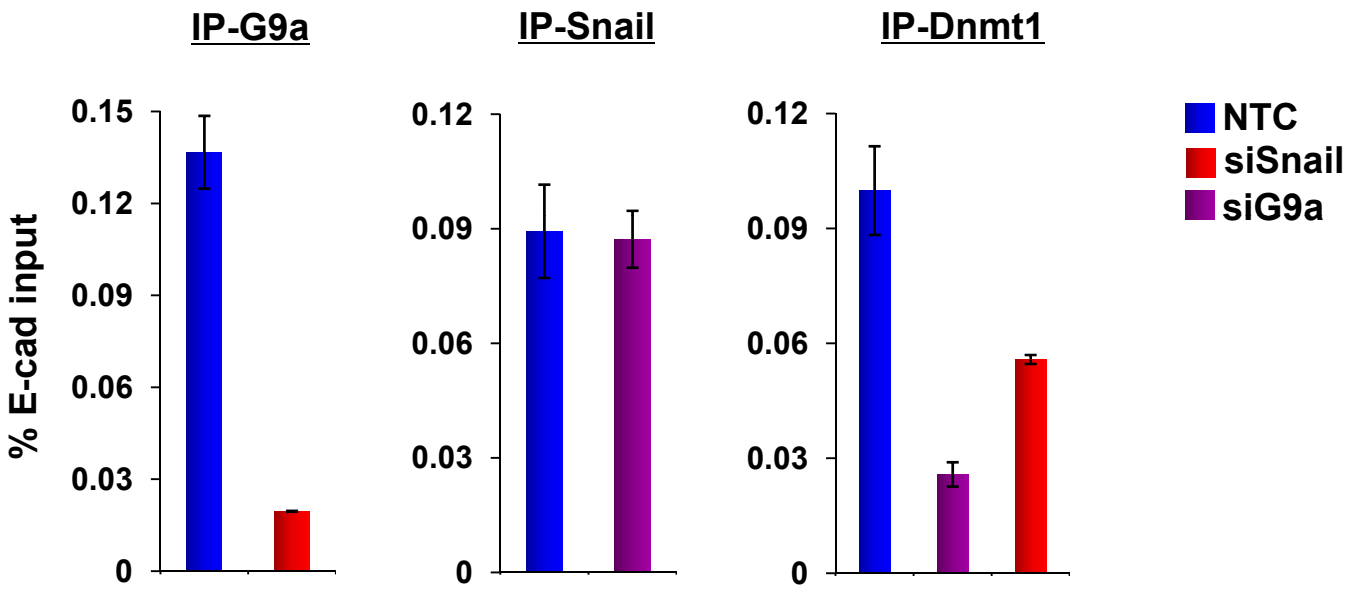


Supplementary Figure S8. G9a-related repressive marks are enriched at the E-cadherin promoter in CLBC cell lines. Quantitative real-time PCR was performed to analyze ChIP samples from Fig. 6B. Results from three independent experiments were presented (mean ± SD from three separate experiments).

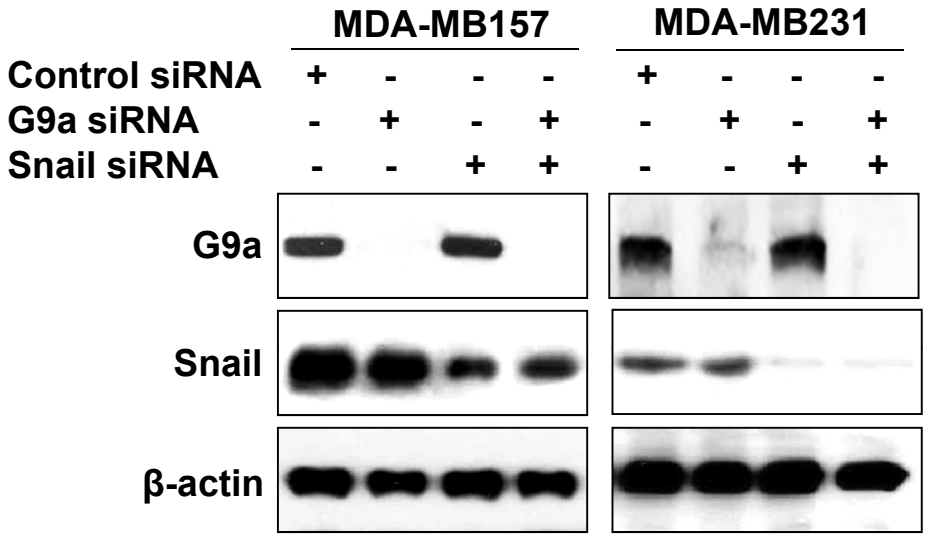




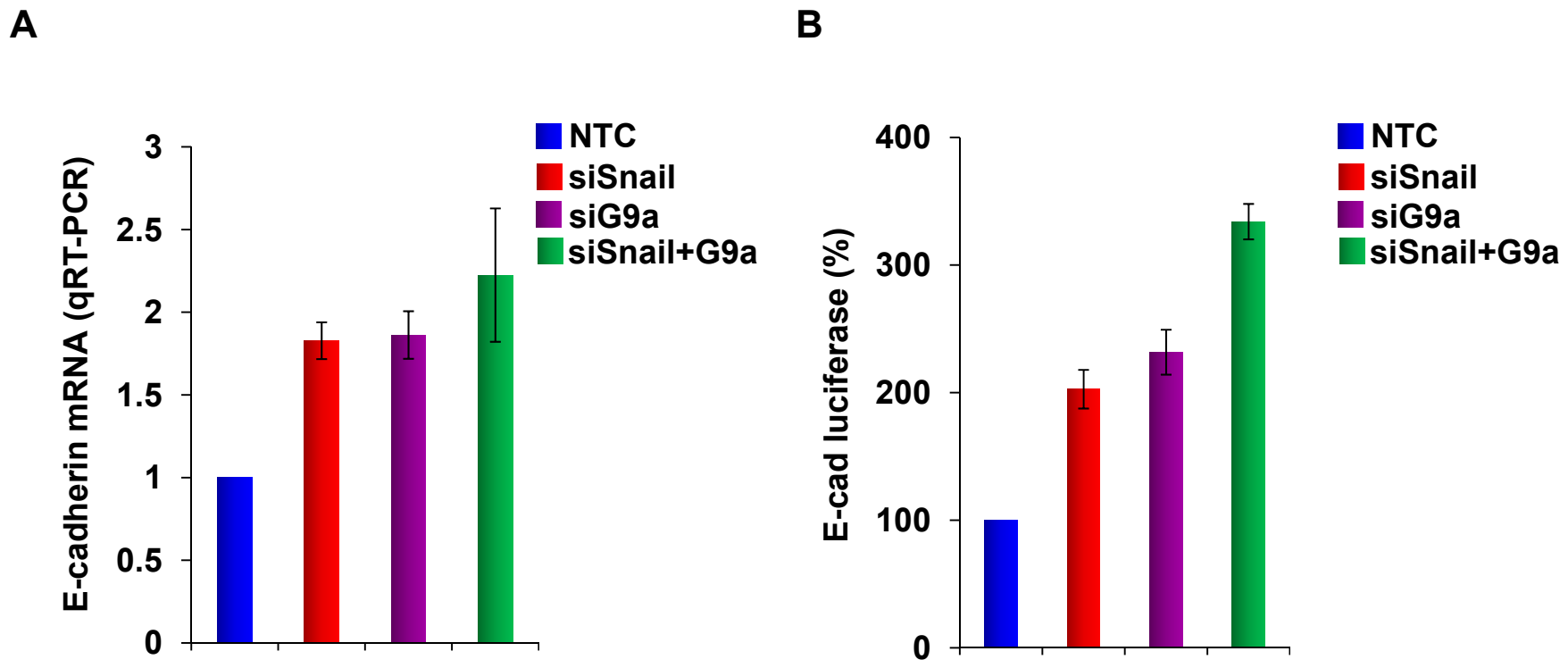
Supplementary Figure S9. Lysates of fresh frozen tumor samples from 25 cases of luminal breast cancer and 16 cases of triple-negative breast cancer were analyzed for the expression of E-cadherin, Claudin 3, ERα, Snail and G9a by Western blotting. Actin was used a loading control.



Supplementary Figure S10. Quantitative real-time PCR was performed to analyze ChIP samples from Fig. 7B. Results from three independent experiments are presented (mean ± SD from three separate experiments).



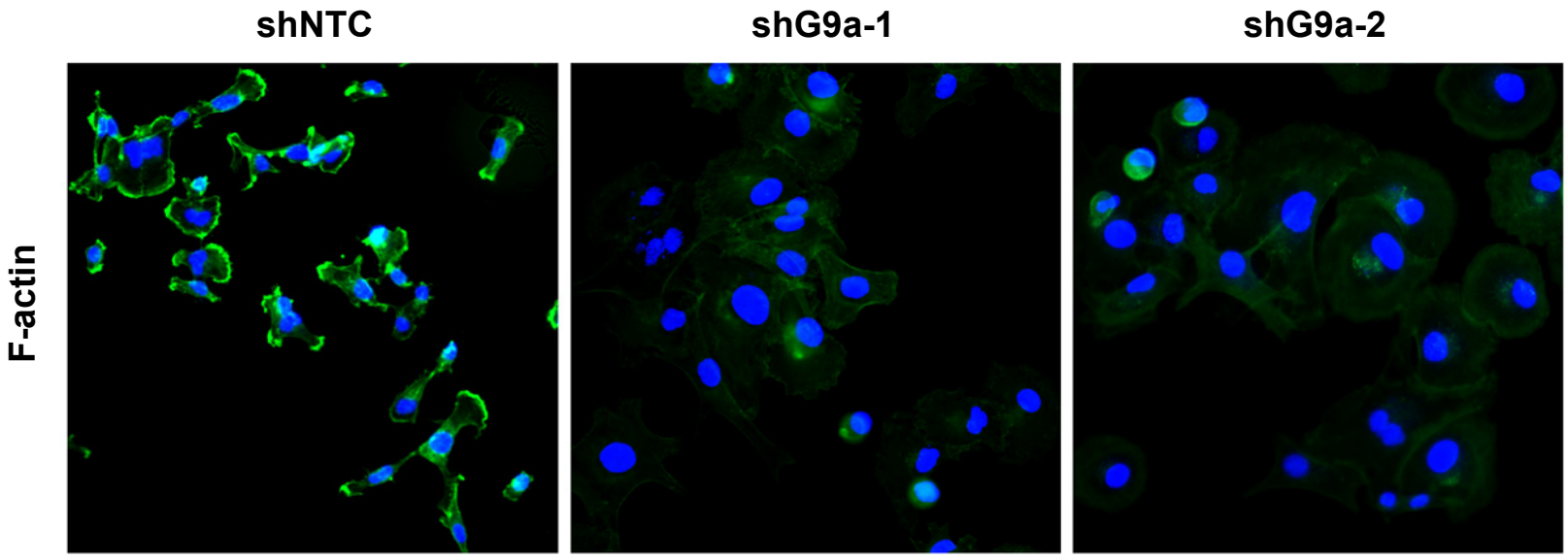
Supplementary Figure S11. Snail, G9a or NTC siRNA was expressed in MDA-MB157 and MDA-MB231 cells. After 48 h, expression of Snail and G9a were analyzed by Western blotting. Actin was used a loading control.



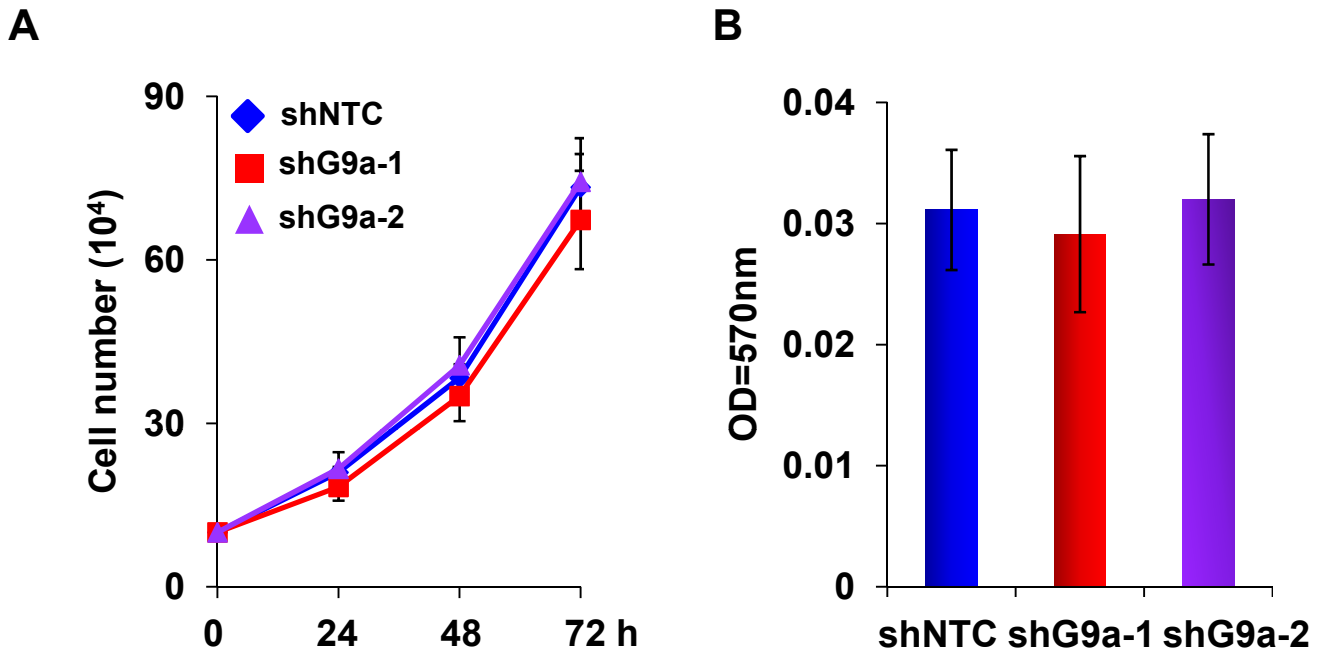
Supplementary Figure S12

(**A**) G9a, Snail or NTC siRNA were expressed in MDA-MB157 cells. After 48 h, E-cadherin mRNA was analyzed by quantitative real-time-PCR.

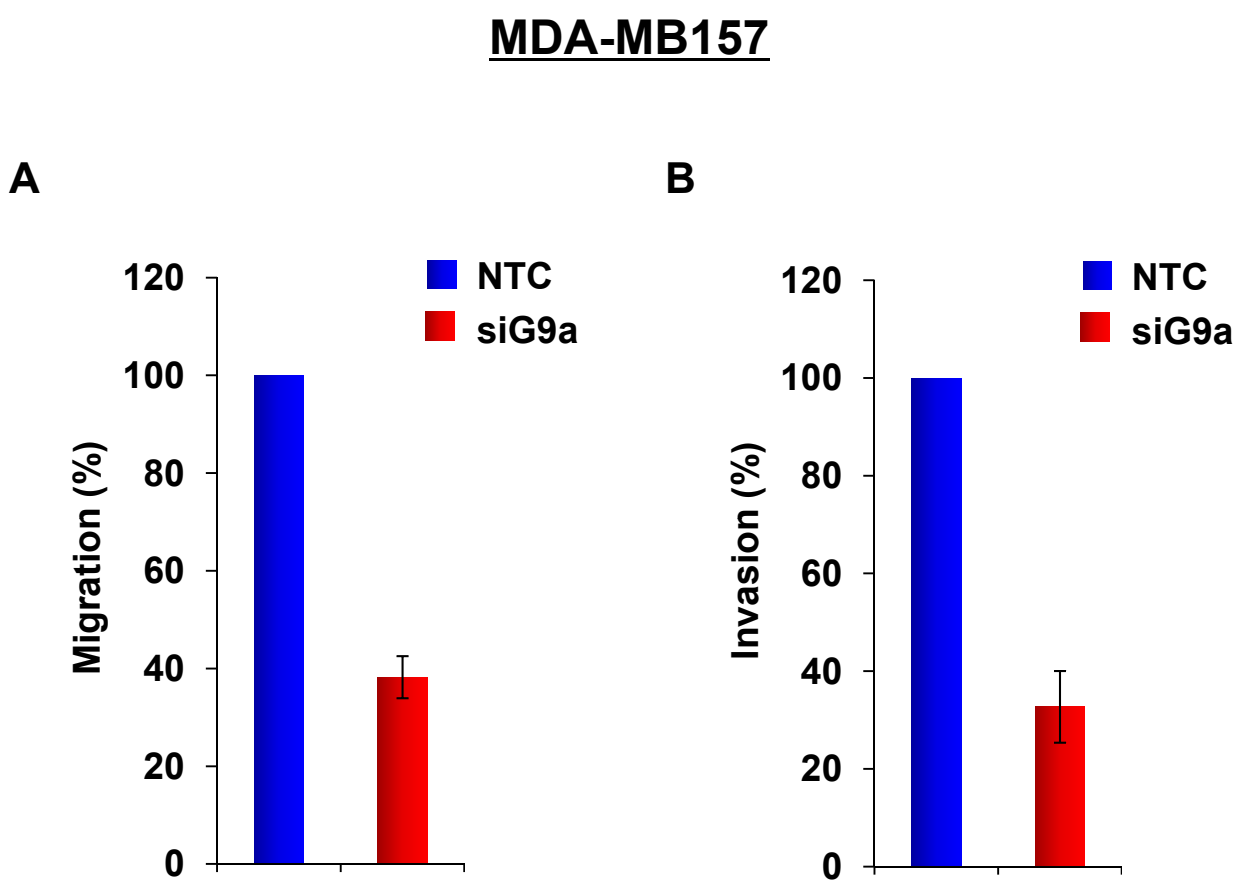
(**B**) G9a, Snail or NTC siRNA were co-expressed with the E-cadherin promoter luciferase construct in MDA-MB157 cells. After 48 h, luciferase activities were normalized and determined (mean ± SD in three separate experiments).



Supplementary Figure S13. Formation of lamellipodia in MDA-MB231 cells and stable transfectants with knockdown of G9a was analyzed by immunofluorescent staining with actin-phalloidin. Nuclei were stained with DAPI (blue).



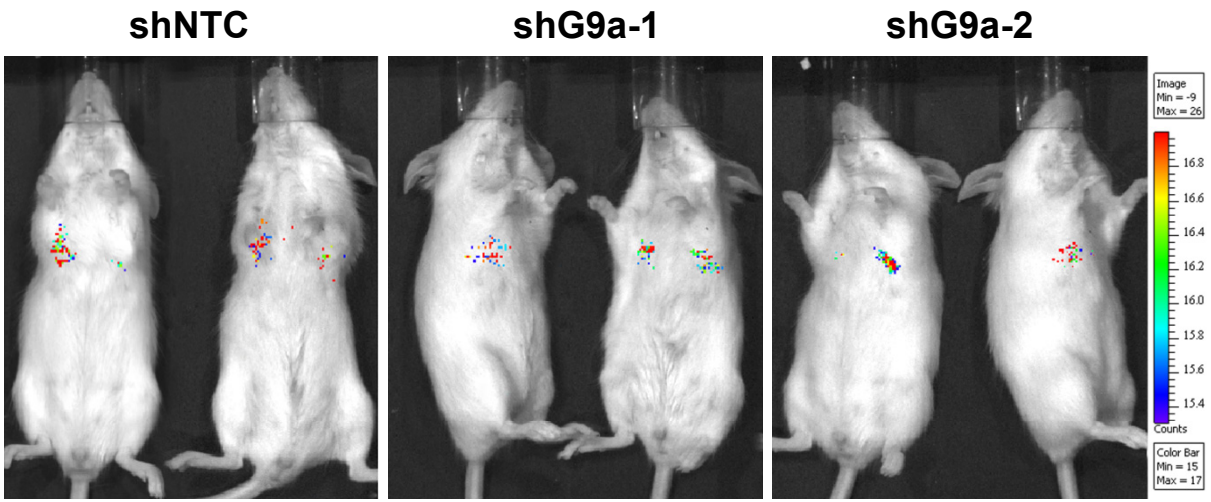
Supplementary Figure S14. Knockdown of G9a expression does not affect the growth and proliferation of MDA-MB231 cells.
(**A**) Cell growth rates of MDA-MB231 cells with stably expressing control vector or G9a shRNA were analyzed by the cell count assay for a period of 3 days. Statistical analyses from two independent experiments with triplicate samples are plotted.
(**B**) Cell proliferation rates of MDA-MB231 cells with stably expressing control vector or G9a shRNA were examined by MTT assays. Statistical analyses from three independent experiments with triplicate samples are shown on the bar graph.

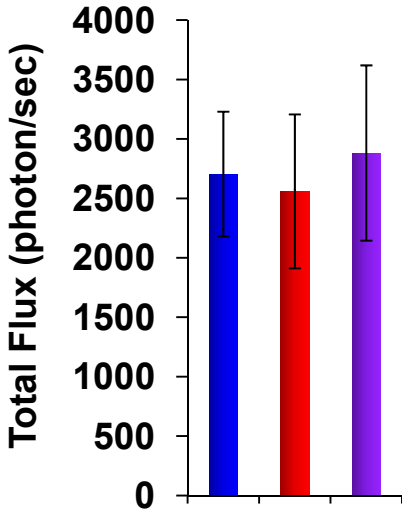


Supplementary Figure S15. Knockdown of G9a expression inhibits migration and invasion of MDA-MB157 cells.

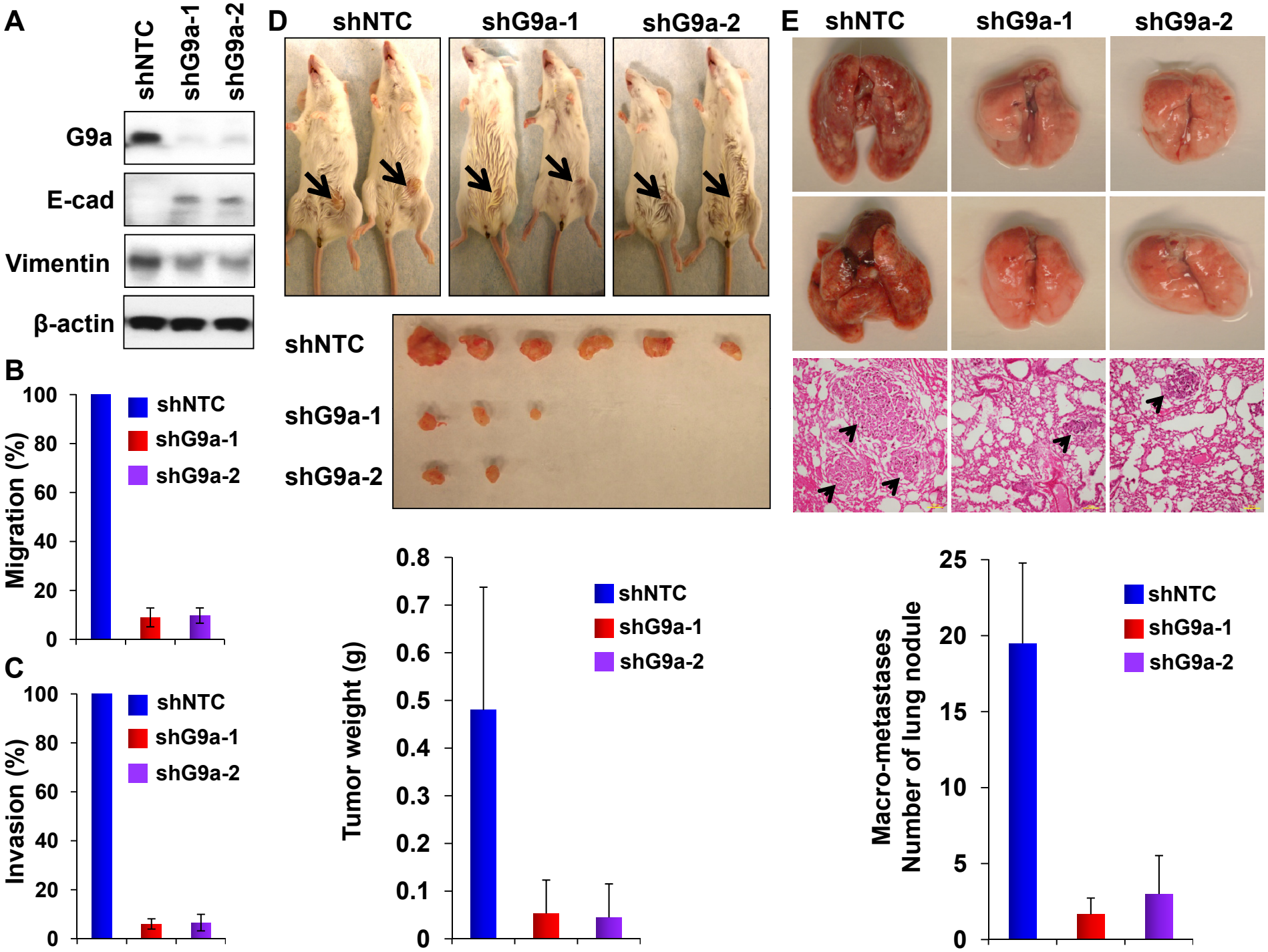
(**A**) The migratory ability of MDA-MB157 cells with and without of knockdown of G9a expression was analyzed by wound healing assay as described in the Materials and Methods. Statistical analysis for the cell migration is shown on the bar graph (mean ± SD from three independent experiments).

(**B**) The invasiveness of MDA-MB157 cells with and without of knockdown of G9a expression was analyzed with a modified Boyden Chamber invasion assay as described in the Materials and Methods. The percentage of invasive cells is shown in the bar graph (mean ± SD in three separate experiments).





Supplementary Figure S16. MDA-MB231 cells with stably expressing control vector or G9a shRNA were injected into the tail vein of SCID mice. To ensure an even number of cell injections from three groups, luciferase bioluminescence imaging measured photon flux after initial injection were taken and quantified.


A
shNTC
shG9a-1
shG9a-2
G9a
E-cad
Vimentin
β-actin
B
Migration (%)
100
80
60
40
20
0
shNTC
shG9a-1
shG9a-2
C
Invasion (%)
100
80
60
40
20
0
shNTC
shG9a-1
shG9a-2
D
shNTC
shG9a-1
shG9a-2
shNTC
shG9a-1
shG9a-2
Tumor weight (g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
shNTC
shG9a-1
shG9a-2
E
shNTC
shG9a-1
shG9a-2
Macro-metastases
Number of lung nodule
25
20
15
10
5
0
shNTC
shG9a-1
shG9a-2

Supplementary Figure S17. Knockdown of G9a expression inhibits breast cancer MDA-MB435 cell migration and invasion *in vitro* and suppresses tumor growth and lung colonization *in vivo*.

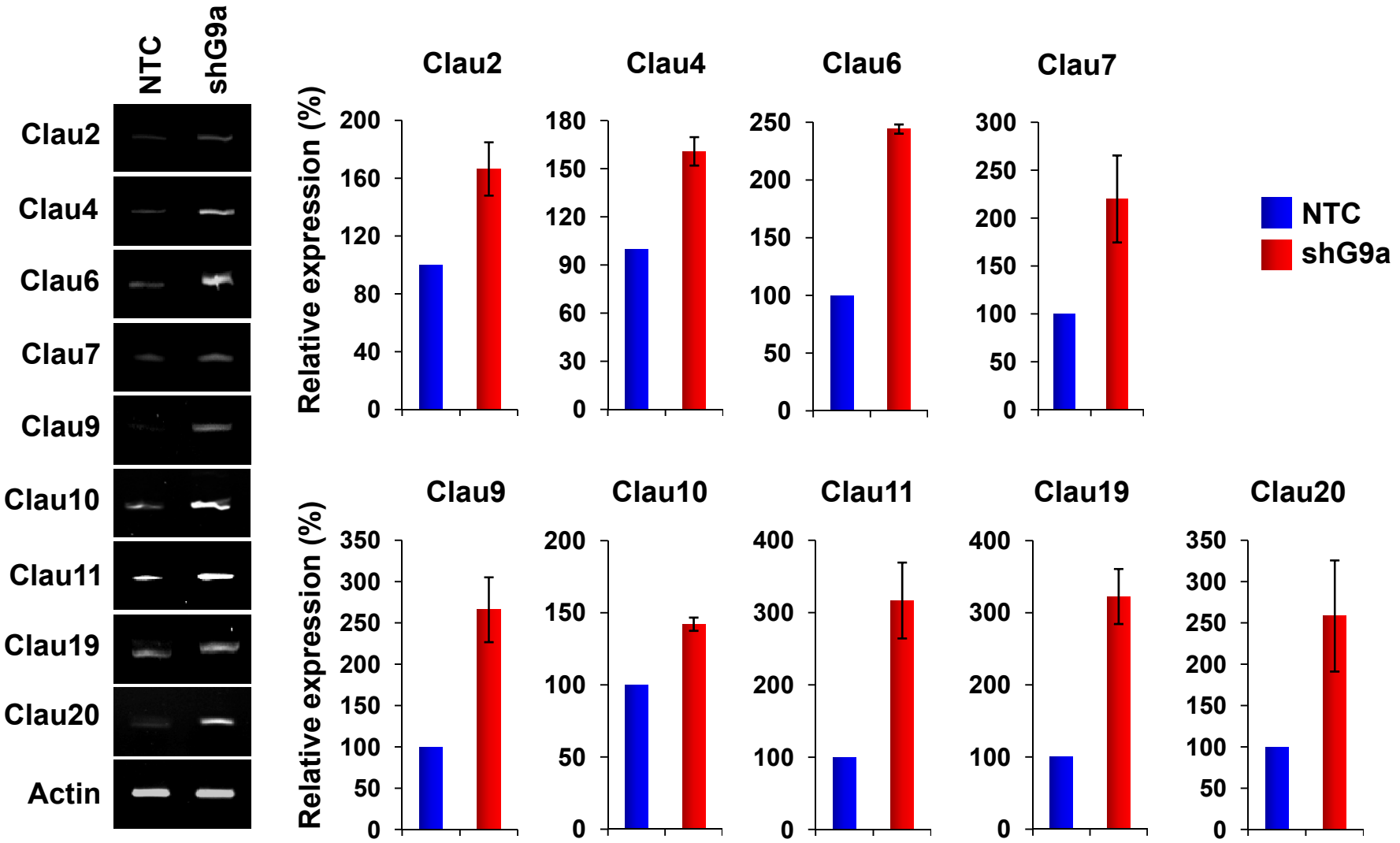
(**A**) MDA-MB435 cells stably expressing control vector or G9a shRNA were examined for the expression of G9a, E-cadherin and vimentin by Western blotting.

(**B**) The migratory ability of MDA-MB435 cells and the corresponding stable transfectants with knockdown of G9a expression were analyzed by wound healing assay as described in Materials and Methods. Statistical analysis for the cell migration is shown in the bar graph (mean $\pm$ SD from three independent experiments).

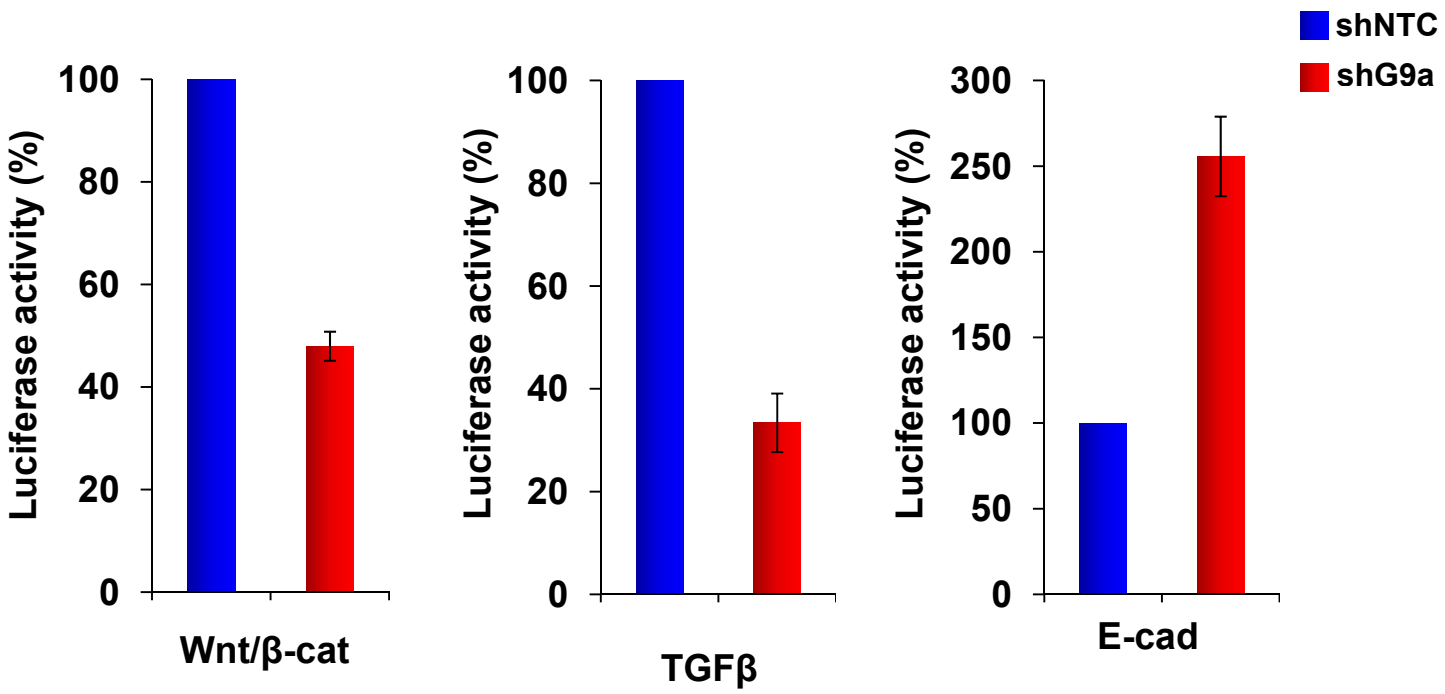
(**C**) The invasiveness of MDA-MB435 cells with stably expressing control vector or G9a shRNA was analyzed with a modified Boyden Chamber invasion assay as described in Materials and Methods. The percentage of invasive cells is shown in the bar graph (mean $\pm$ SD in three separate experiments).

(**D**) MDA-MB435 cells with stably expressing control vector or G9a shRNA were injected into the mammary fat pad of SCID mice. The growth of breast tumors was measured weekly by caliber. Mice were sacrificed when tumor of control group reached 1.5 cm. Tumors were removed and weighed (the mean of 6 animals $\pm$ SEM).

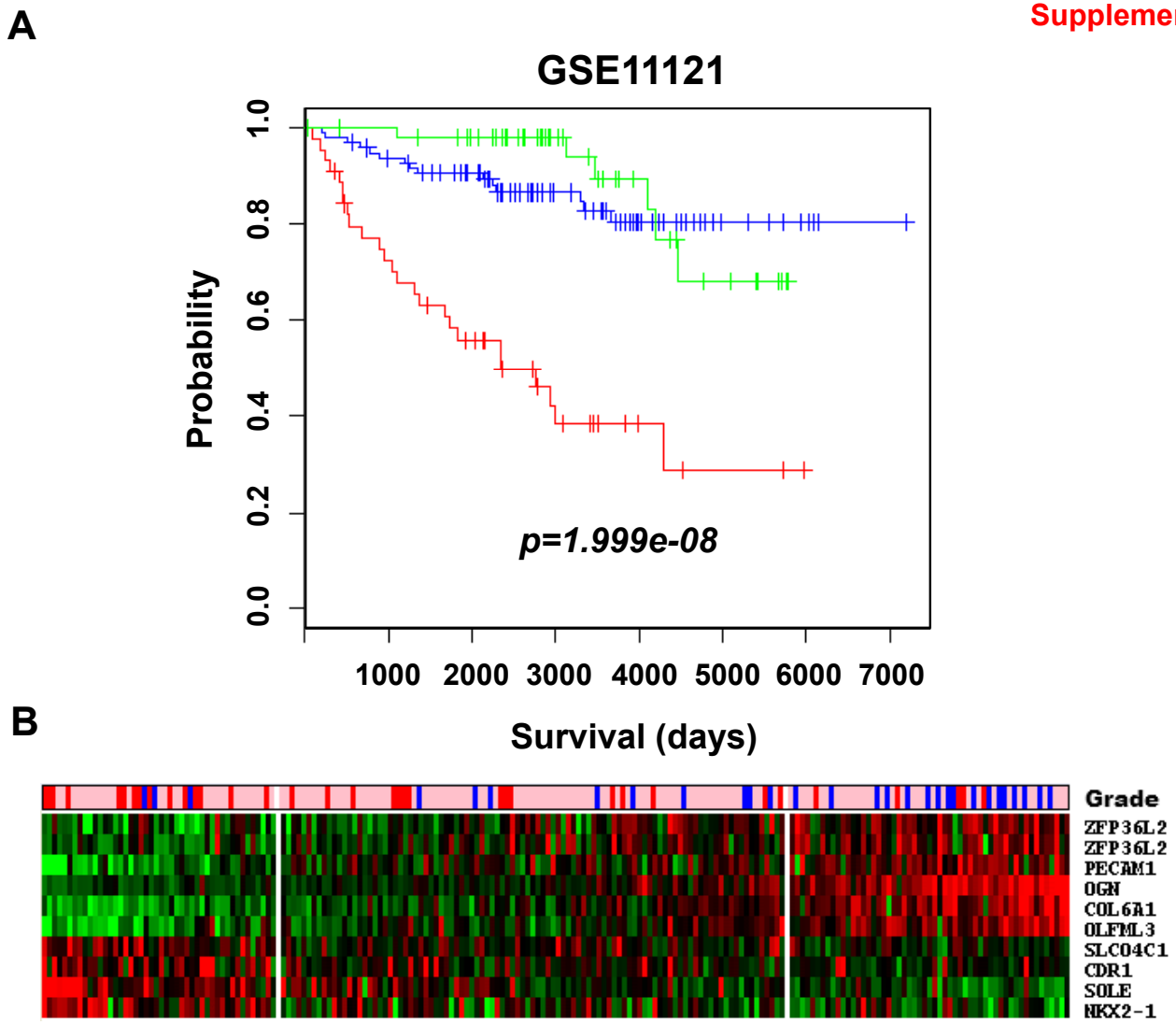
(**E**) MDA-MB435 cells with stably expressing control vector or G9a shRNA were injected into the tail vein of SCID mice. Mice were sacrificed after 12 weeks; lung metastatic nodules were examined macroscopically or detected in paraffin-embedded sections stained with hematoxylin and eosin. The number of macro-metastatic nodules was presented (the mean of 6 animals $\pm$ SEM).



Supplementary Figure S18. Knockdown of G9a restored claudin expression in MDA-MB231 cells. The differentially expressed claudins from MDA-MB231 vector-control cells and the corresponding stable clone with knockdown of G9a expression were analyzed by either quantitative RT-PCR (left panel) or real-time PCR. Results from three independent experiments of real-time PCR were presented on the right panel (mean ± SD from three separate experiments).



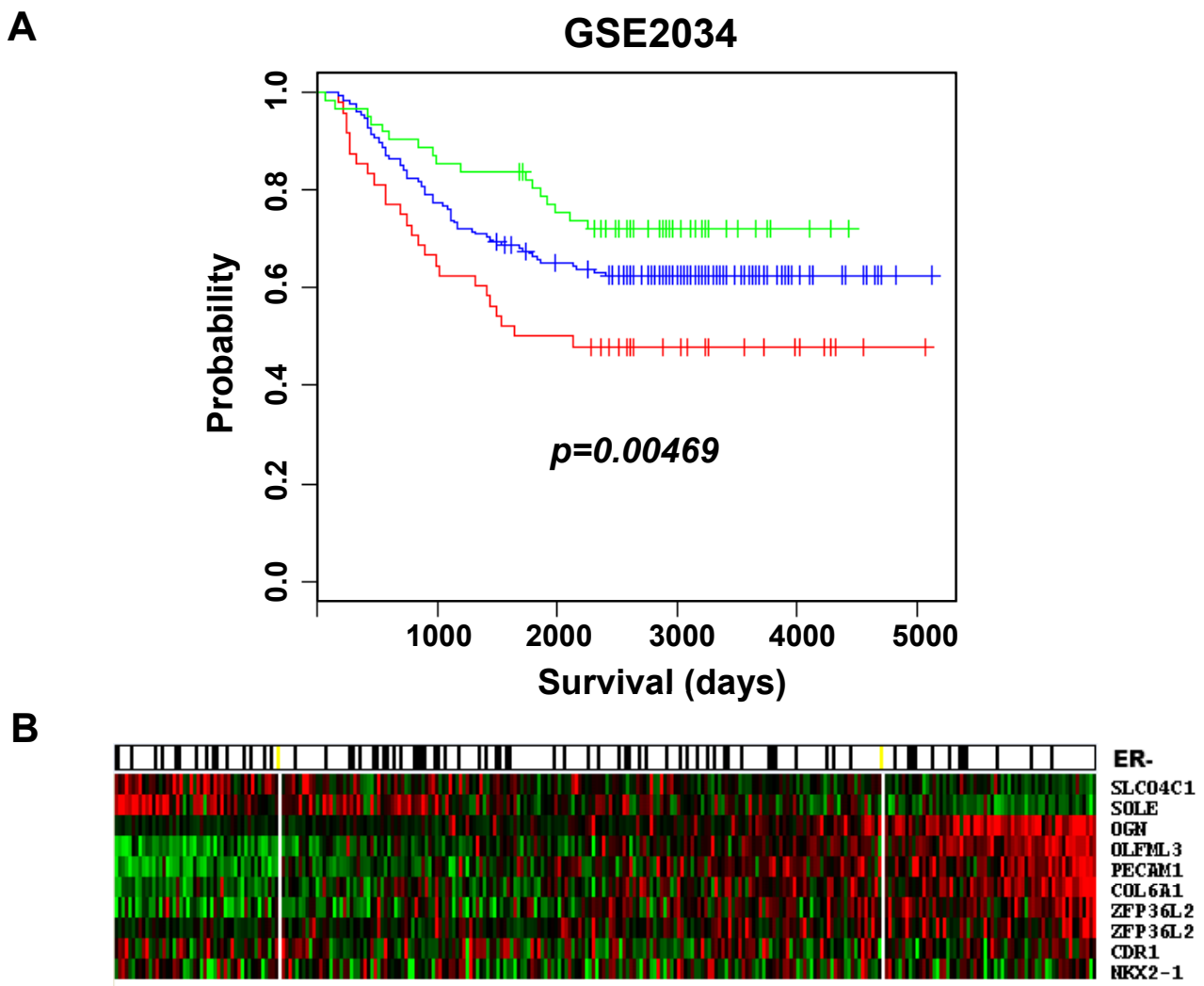
Supplementary Figure S19. Activation of the β-catenin, TGF-β and cellular adhesion pathways in MDA-MB231 cells and corresponding stable transfectants with knockdown of G9a expression was analyzed by luciferase reporter assays (mean ± SD in three separate experiments).



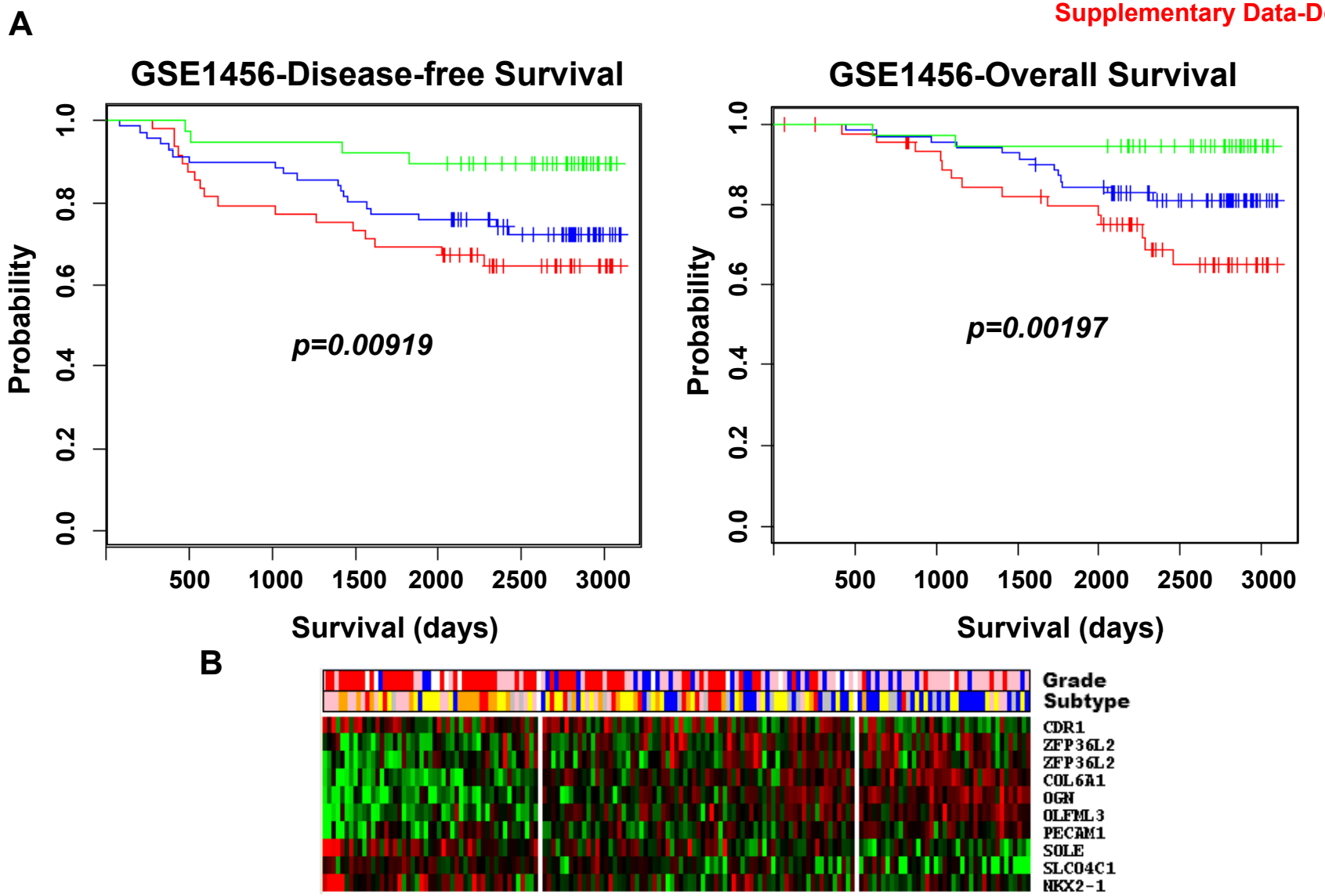
Supplementary Figure S20. Identification of a 9-gene prognostic signature using breast cancer dataset GSE11121.

(**A**) Kaplan–Meier survival curve separates the tumors (from GSE11121 data set, training set) into three groups with expression of a 9-gene prognostic signature.

(**B**) Heatmap of the 9-gene signature expression in 200 lymph node-negative breast cancer patients. Top bar, tumor grade (1=blue, 2=pink, 3=red).



Supplementary Figure S21. Validation of the 9-gene prognostic signature using breast cancer dataset GSE2034.
(**A**) Kaplan-Meier plot of patient survival as stratified using the 9-gene signature.
(**B**) Heatmap of the 9-gene signature expression in 286 lymph node-negative breast cancer patients. Top bar: ER status (white, ER positive; black, ER negative).



Supplementary Figure S22. Validation of the 9-gene prognostic signature using breast cancer dataset GSE1456.
(**A**) Kaplan-Meier plot of disease-free and overall patient survival as stratified using the 9-gene signature.
(**B**) Heatmap of the 9-gene signature expression in 159 breast cancer patients. Top bars: tumor grade (1=blue, 2=pink, 3=red) and tumor subtypes (normal-like, blue, luminal-A yellow, luminal-B orange, basal, pink, HER2, red).