Supplemental Data

Materials and Methods

Quantatitive PCR. Total RNA isolated from FACS-sorted subpopulations based upon the expression of CD29 and CD24 were extracted using the RNA purification kit as mentioned in "Microarray Analysis". RNA (300 ng) each was used to generate cDNA with the High Capacity cDNA Reverse Transcription Kit. The DNA concentration was measured using a spectrophotometer (NanoDrop), and 1,000 pg of DNA generated from the above reaction was applied in Taqman Gene Expression Master Mix for quantitative PCR. Quantatitive PCR was performed on the StepOnePlus Real-time PCR System with the following steps: 50°C for 2 min, 95°C for 10 min, and then 40 cycles for 95°C 15 sec and 60°C 1 min. 18S rRNA was used to normalize and calculate relative gene expression. All reagents and instruments, if not mentioned, including six probe and primer sets were from Applied Biosystems.

Table S1 Histopathology of the p53 null-mammary gland tumors used in this study and their corresponding ERα status

Paraffin-embedded slides were stained with anti-K5, K14, K8, and ERα antibodies, and reviewed by an expert pathologist. Primary tumors and their corresponding transplantation generations display very similar histology. *T1, T2, and T7 tumors were used for RNA isolation, and subsequent microarray analysis.

Table S1

Tumor	Tumor Type	ERα Status
T1*	Squamous adenocarcinoma, heavily keratinizing	Negative
T2*	Typical papillary adenocarcinoma, well differentiated	Positive
T3	Papillary adenocarcinoma	Positive
T4	Solid adenocarcinoma, usual undifferentiated type	Positive
T5	Very mixed and undifferentiated carcinoma[1]	Positive
T6	Typical solid adenocarcinoma, undifferentiated	Negative
T7*	Solid adenocarcinoma, usual undifferentiated type	Negative
T8	Solid adenocarcinoma	Negative
T9	Solid adenocarcinoma	Negative
T10	Typical solid adenocarcinoma, undifferentiated	Negative

[1]areas of proliferation of both luminal cells and myoepithelial cells

*tumors used in microarray analysis

Table S2 The $Lin^{-}CD44^{+}CD24^{-}$ subpopulation isolated from p53 null-mammary tumors showed a similar tumor initiation potential as other subpopulations

Cells from T2, T3, T9 (all TG 1), were freshly digested and FACS sorted. All dead cells and lineage positive cells were gated out. Designated number of cells were washed with PBS and transplanted into cleared fat pads of 3-week-old Balb/c female mice. Mice were monitored until tumors were observed or up to 18 months if no tumors were detected. Tumor-initiating cell frequency and the Poisson distribution analysis was generated using R software (The R Foundation for Statistical Computing).

Table S2

	5,000	2,000	1,000	500	100	Tumor-initiating cell frequency (95% CI)
$Lin^-CD44^+CD24^-$	7/12	3/5	4/6	1/6	0/10	1/3524 (1/6153 - 1/2018)
$Lin^-CD44^+CD24^+$	8/12	3/5	5/6	1/6	0/10	1/2832 (1/4854 - 1/1652)
$Lin^-CD44^-CD24^+$	6/12	1/5	3/6	1/6	0/4	1/5386 (1/10039 - 1/2890)
$Lin^-CD44^-CD24^-$	5/12	0/5	3/6	0/6	0/2	1/7984 (1/16253 - 1/3922)

Table S3 Tumor-initiating cells were enriched in $Lin^-CD29^HCD24^H$ subpopulation of individual p53 null mammary tumors. Tumor-initiating cell frequency and the Poisson distribution analysis was generated using R software (The R Foundation for Statistical Computing) where applicable.

Table S3

Tumor T1	30,000	5,000		1,500	500	100	Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		3/3		2/2	5/5	2/3	1/87 (1/310 - 1/24)
$Lin^-CD29^HCD24^L$		3/3		1/2	2/5	0/2	1/1312 (1/3532 - 1/487)
$Lin^-CD29^LCD24^H$		1/4		0/2	1/3	0/1	1/10828 (1/45639 - 1/2569)
$Lin^-CD29^LCD24^L$		1/4		0/2	0/3	0/1	1/22005 (1/154231 - 1/3140)
Lin^-	1/1	1/1		0/2			1/5015 (1/30358 - 1/828)

Tumor T2	30,000	5,000		1,500	500	100	Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		3/3		2/2	4/5	1/3	1/291 (1/742 - 1/114)
$Lin^-CD29^HCD24^L$		3/4		1/2	2/5	0/2	1/2404 (1/6129 - 1/943)
$Lin^-CD29^LCD24^H$		1/4		0/2	0/3	0/2	1/22105 (1/154881 - 1/3155)
$Lin^-CD29^LCD24^L$		0/4		0/2	0/3	0/2	
Lin^-	1/1	0/2		0/2			1/25078 (1/162992 - 1/3858)

Tumor T3	30,000	5,000			500		Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		3/3			4/5		1/311 (1/923 - 1/104)
$Lin^-CD29^HCD24^L$		3/4			1/5		1/3218 (1/9457 - 1/1095)
$Lin^-CD29^LCD24^H$		1/4					
$Lin^-CD29^LCD24^L$		0/4					
Lin^-	0/1	0/2					

Tumor T4	30,000	5,000	2,500		500		Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		4/4			4/5		1/311 (1/1923 - 1/105)
$Lin^-CD29^HCD24^L$		3/4			2/5		1/2412 (1/6975 - 1/834)
$Lin^-CD29^LCD24^H$		1/6	0/2				1/32436 (1/229416 - 1/4586)
$Lin^-CD29^LCD24^L$		1/6	0/2				1/32436 (1/229416 - 1/4586)
Lin^-	1/1	0/2	1/2				1/11386 (1/52888 - 1/2451)

Tumor T5	30,000	5,000			500		Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		4/4			5/5		
$Lin^-CD29^HCD24^L$		3/4			2/5		1/2412 (1/6975 - 1/834)
$Lin^-CD29^LCD24^H$		1/6					
$Lin^-CD29^LCD24^L$		0/6					
Lin^-	1/1	0/1					1/15417 (1/120906 -1/1966)

Tumor T6	30,000	5,000	2,500	1,500	500	100	Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		4/4		1/2	4/6	2/2	1/556 (1/1368 - 1/226)
$Lin^-CD29^HCD24^L$		4/5		1/2	2/6	0/2	1/2367 (1/5592 - 1/1002)
$Lin^-CD29^LCD24^H$		1/4	1/2	0/2	0/2	0/1	1/12571 (1/50879 - 1/3106)
$Lin^-CD29^LCD24^L$		0/4	1/2	0/2	0/2	0/1	1/27831 (1/203229 - 1/3811)
Lin^-	1/1	1/2	0/2	0/2			1/12401 (1/53063 - 1/2898)

Tumor T7	30,000	5,000		1,500	500	100	Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		4/4		2/2	3/5	2/3	1/346 (1/875 -1/137)
$Lin^-CD29^HCD24^L$		4/4		0/2	1/6	0/3	1/2604 (1/6533 - 1/1038)
$Lin^-CD29^LCD24^H$		1/4		0/2	0/3	0/1	1/22005 (1/154231 - 1/3140)
$Lin^-CD29^LCD24^L$		0/4		0/2	0/3	0/1	
Lin^-	1/1	1/2		0/2			1/7629 (1/40867 - 1/1424)

Tumor T8	30,000	5,000			500		Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		4/4			5/5		
$Lin^-CD29^HCD24^L$		4/4			1/6		1/1674 (1/4822 - 1/581)
$Lin^-CD29^LCD24^H$		0/4					
$Lin^-CD29^LCD24^L$		0/4					
Lin^-	0/1	0/2					

Tumor T9	30,000	5,000			500		Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		3/3			5/5		
$Lin^-CD29^HCD24^L$		4/4			1/5		1/1499 (1/4606 - 1/488)
$Lin^-CD29^LCD24^H$		0/4					
$Lin^-CD29^LCD24^L$		0/4					
Lin^-	1/1	0/2					1/21639 (1/146344 - 1/3199)

Tumor T10	30,000	5,000		1,500	500	100	Tumor-initiating cell frequency (95% CI)
$Lin^-CD29^HCD24^H$		4/4		2/2	5/6	1/3	1/270 (1/649 - 1/112)
$Lin^-CD29^HCD24^L$		4/4		1/3	2/6	0/3	1/1740 (1/4089 - 1/740)
$Lin^-CD29^LCD24^H$		0/4		0/2	0/3	0/1	
$Lin^-CD29^LCD24^L$		0/4		0/2	0/2	0/1	
Lin^-	1/1	0/2		0/2			1/25078 (1/162992 - 1/3858)

Table S4 The $Lin^-CD49f^HCD24^H$ subpopulation from p53 null mammary tumor showed increased tumor initiation potential as compared with all other Lin-CD49fCD24 subpopulations

Cells from T7 (primary tumor), T2, T9 (TG 1), were freshly digested and FACS sorted. All dead cells and lineage positive cells were gated out. Designated number of cells were washed with PBS and transplanted into cleared fat pad of 3-week-old Balb/c female mice. Mice were monitored until tumors were observed or up to 18 months if no tumors were detected. Tumor-initiating cell frequency and the Poisson distribution analysis was generated using R software (The R Foundation for Statistical Computing).

Table S4

	5,000	2,000	1,000	500	100	Tumor-initiating cell frequency (95% CI)
$Lin^-CD49f^HCD24^H$	4/4	5/6	3/3	3/6	1/4	1/863 (1/1673 – 1/445)
$Lin^-CD49f^HCD24^L$	0/3	1/6	0/3	0/6		1/31989 (1/231178 – 1/4426)
$Lin^-CD49f^LCD24^H$	1/9	1/6	0/3			1/28107 (1/113851 – 1/6988)
$Lin^-CD49f^LCD24^L$	0/6	0/6	0/3			

Table S5 List of genes differentially expressed in $Lin^-CD29^HCD24^H$ vs. other subpopulations from p53 null-mammary tumor cells, and genes differentially expressed in normal $Lin^-CD29^HCD24^P$(MEC) vs. $Lin^-CD29^PCD24^H$ (MEC) and Lin^-CD24^L(MEC). $p<0.01$

Table S5

Attached Excel file

Figure Legends

Figure S1 FACS profiles comparison between serial passaged p53 null mammary tumors. A. A p53 null mammary gland tumor, T2, B. The first transplantation generation of tumor T2 (TG1), and C. The second transplantation generation of tumor T2 (TG2), were analyzed by FACS based upon the expression of the cell surface markers, CD29 and CD24.

Figure S2 Flow analysis of p53 null mammary tumors based upon the expression of the human tumor- initiating markers, CD44 and C24. A: Tumor T2, B. Tumor T3.

Figure S3 FACS profile comparison between the parental tumor and the subsequent $CD29^HCD24^H$ subpopulation-derived tumor, $CD29^HCD24^L$ subpopulation-derived tumor. A. A p53 null mammary gland tumor, T6, was FACS sorted based upon expression of CD29 and CD24. Individual subpopulations were collected and transplanted as described in Materials and Methods under “Transplantation into the Clear

Fat Pad". Tumors generated from 1,500 cells of $Lin^-CD29^HCD24^H$ (B), and 1,500 cells of $Lin^-CD29^HCD24^L$ (C) were further analyzed based upon CD29, CD24 expression.

Figure S4 The $Lin^-CD29^HCD24^H$ subpopulation contains more K8/K14 double positive cells than other $Lin^-CD29CD24$ subpopulations. A. FACS sorted $Lin^-CD29^HCD24^H$ (a, c) and other $Lin^-CD29CD24$ cells (b, d) from p53-null mammary tumors, T2 (a, b) and T7 (c, d) were co-stained with anti-K8 antibody (red), anti-K14 antibody (green) and DAPI (blue). Scale bar, 25 μm. B. The data were analyzed based upon co-staining data from tumors T1, T2, and T7. Two biological replicates with a total of 3,000 cells from each tumor were counted. $P<0.0007$.

Figure S5 Flow analysis of p53 null mammary tumors based upon CD49f and CD24 expression. FACS profiling of T2 (A) and T9 (B) tumors using the mouse mammary gland stem cell markers, CD49f and CD24, (a, d); histograms of FITC-isotype controls and the corresponding CD49f-FITC antibodies (b, e); histograms of R-PE-isotype controls and the corresponding CD24-PE antibodies (c, f).

Figure S6 Mammosphere formation from single cells. Primary mammospheres from tumor T7 were dissociated with trypsin as mentioned in Materials and Methods under "Mammosphere assays". Single cells were plated in 96-well low attachment plates. One thousand wells each containing single cells were monitored and later counted. Pictures were taken under 3.2X magnification. The efficiency of mammosphere formation was 7%.

Figure S7 p53 null mammospheres are enriched in the $Lin^-CD29^HCD24^H$ tumor-initiating cells. A. Tertiary mammospheres from tumor T7 were dissociated with trypsin, then labeled and FACS sorted based upon the expression of CD29 and CD24. B. Collagenase digested cells from tumor T7 were plated on plastic with mammosphere medium supplemented with 5% FBS. Cells were collected with trypsin, then labeled and FACS sorted based upon the expression of CD29 and CD24. Experiments done on tumor T3 showed the similar results.

Figure S8 Quantitative PCR (qPCR) validation of microarray gene expression. A. The fold change in gene expression determined from microarray analysis (A, T2; C, T7) was calculated using dChip PM (Perfect-Match) model with quantile normalization as described in Materials and Methods under "Microarray Analysis". B. Six probes Prom1 (Prominin 1), Ezh2 (Enhancer of Zeste homolog 2), Ect2 (Ect2 oncogene), Bub1 (budding uninhibited by benzimidazoles 1 homolog), Hus1 (HUS1 checkpoint homolog (S. pombe)), and Itga6 (Integrin alpha 6, CD49f) (all purchased from Applied Biosystems) were used in quantatitive PCR analysis to validate the microarray data (B, T2; D, T7). In both array and qPCR analysis, the fold change of the subpopulation with the lowest expression was designated as "1" for each probe. Three biological replicates were included in each sample. ▒ $Lin^-CD29^HCD24^H$, ■ $Lin^-CD29^HCD24^L$, □ $Lin^-CD29^LCD24^H$, ▓ $Lin^-CD29^LCD24^L$.

Figure S9 Mammary epithelial cells (MEC) from 8 to 10-week-old Balb/c female mice were flow sorted based upon the expression of CD29 and CD24. A. The percentage of each individual subpopulation was generated based upon the negative isotype controls. B and C. Histogram plots of MEC shown with arrows indicating the isotype controls and their corresponding antibodies.