Supplementary Material

Supplementary tables S1-S6,

Supplementary methods,

Supplementary figures S1-S4.

Supplementary tables S1-S6

Supplementary table S1 Tumorigenic abilities of 20th passage BCSC and their primary breast cancer cells in NOD/SCID mice.

	Tumors/injections			
	1×10^2	1×10^3	1×10^4	1×10^5
Cell line 1				
MDA-MB-231(unsorted)	0/8	0/8	0/8	0/8
MDA-MB-231.SC	0/8	5/8*	8/8**	—
Cell line 1				
MDA-MB-231($CD24^-CD44^+$)	0/8	4/8	8/8	—
MDA-MB-231.SC	0/8	5/8	8/8	—
Mouse passage 1 of MDA-MB-231.SC				
Unsorted	—	0/8	3/8	8/8
$CD24^-CD44^+$	—	6/8**	8/8*	—
Mouse passage 2 of MDA-MB-231.SC				
$CD24^+CD44^+$	—	—	0/8	—
$CD24^-CD44^+$	—	—	7/8**	—
Cell line 2				
MCF-7(unsorted)	0/8	0/8	0/8	0/8
MCF-7.SC	0/8	5/8*	7/8**	—
Cell line 2				
MCF-7($CD24^-CD44^+$)	0/8	4/8	7/8	—
MCF-7.SC	0/8	5/8	7/8	—
Mouse passage 1 of MCF-7.SC				
Unsorted	—	0/8	3/8	8/8
$CD24^-CD44^+$	—	7/8**	8/8*	—
Mouse passage 2 of MCF-7.SC				
$CD24^+CD44^+$	—	—	0/8	—
$CD24^-CD44^+$	—	—	8/8***	—

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Supplementary table S2 MiroRNA expressions were regulated simultaneously in both BC (breast cancer cells) and BCSC (breast cancer stem cells) in comparison with immortalized healthy mammary epithelial cell lines (HME).

Downregulated miRNAs in BC and BCSC

Name of miRNA	Mature sequence	Fold change (HME/BC)	Fold change (HME/BCSC)
hsa-miR-34a-5p	ACAACCAGCTAAGACACTGC	14.805	14.175
hsa-let-7b-5p	AACCACACAACCTACTACC	2.415	2.207
hsa-miR-200c-3p	TCCATCATTACCCGG	2.719	2.850
hsa-miR-224-5p	AACGGAACCACTAGTGACTT	2.435	2.366
hsa-miR-3663-3p	GCGCCCGGCCT	2.008	2.650

Upregulated miRNAs in BC and BCSC

Name of miRNA	Mature sequence	Fold change (BC/HME)	Fold change (BCSC/HME)
hsa-miR-103a-3p	TCATAGCCCTGTACAATG	4.692	7.096
hsa-miR-107	TGATAGCCCTGTACAATGCT	4.127	5.162
hsa-miR-1234-5p	CGGCCCCCCCC	2.857	5.495
hsa-miR-125b-5p	TCACAAGTTAGGGTCTC	4.928	32.147
hsa-miR-151a-5p	ACTAGACTGTGAGCTCC	2.138	3.009
hsa-miR-15b-5p	TGTAAACCATGATGTGCTGC	4.017	6.126
hsa-miR-16-5p	CGCCAATATTTACGTGCTG	4.448	2.861
hsa-miR-19b-3p	TCAGTTTTGCATGGATTTGC	3.308	2.610
hsa-miR-20a-5p	CTACCTGCACTATAAGCAC	3.185	2.927
hsa-miR-21-5p	TCAACATCAGTCTGATAAGC	2.042	4.816
hsa-miR-23b-3p	GGTAATCCCTGGCAATG	7.423	15.619
hsa-miR-24-3p	CTGTTCCTGCTGAACTGA	3.226	5.175
hsa-miR-26b-5p	ACCTATCCTGAATTACTTGA	7.368	8.879
hsa-miR-27a-3p	GCGGAACTTAGCCACTG	3.011	2.018
hsa-miR-27b-3p	GCAGAACTTAGCCACTGT	4.809	11.652
hsa-miR-29a-3p	TAACCGATTTCAGATGGTGC	2.276	4.239
hsa-miR-29b-3p	AACACTGATTTCAAATGGTGC	3.385	7.237
hsa-miR-30a-5p	CTTCCAGTCGAGGATG	12.057	2.941
hsa-miR-365a-3p	ATAAGGATTTTTAGGGGCATTA	2.936	3.890
hsa-miR-4443	AAAACCCACGCCTCC	6.377	12.019
hsa-miR-4459	CTCCACCTCCTCCG	3.027	3.115
hsa-miR-4530	CGCTCCCGTCCTG	3.011	4.184
hsa-miR-494	GAGGTTTCCCGTGTA	2.489	3.244
hsa-miR-6087	GCTCGCCCCCCC	2.486	4.438
hsa-miR-6088	CGCCCCCCCGC	4.666	10.032

Supplementary table S3 Clinicopathological characteristics of 134 patients and their associations to miR-34a expression by qRT-PCR.

	No. of cases	Low expression of miR-34a (< median)	High expression of miR-34a (> median)	p value
Totality		67	67	
Age (years)				0.076
<50	82	36	46	
≥50	52	31	21	
Pathogenetic location				0.604
Left breast	65	31	34	
Right breast	69	36	33	
Family history of cancer				0.731
Absence	125	62	63	
Presence	9	5	4	
Histological type				0.999
Ductal	128	64	64	
Lobular	6	3	3	
Tumor size (cm)				0.701
≤2	38	18	20	
>2	96	49	47	
Positive axillary nodes				0.446
0	54	24	30	
1-3	33	15	18	
4-9	22	13	9	
≥10	25	15	10	
Pathological staging				<0.001*
I-II	74	23	51	
III	60	44	16	
Pathological grading				0.767
1	34	19	15	
2	39	18	21	
3	51	26	25	
Undifferentiated	9	4	5	
Not available	1	0	1	
ER status				0.021*
Negative	83	48	35	
Positive	51	19	32	
PR status				0.035*
Negative	80	46	34	
Positive	54	21	33	
HER2 status				0.005*
Negative	112	62	50	
Positive	22	5	17	
Triple-negative breast cancer				0.001*
Absence	83	28	55	
Presence	51	39	12	
Local relapse				0.145
Absence	126	61	65	
Presence	8	6	2	
Distant metastatic relapse				0.002*
Absence	98	41	57	
Presence	36	26	10	
p53 status				0.001*
Negative	57	38	19	
Positive	77	29	48	

* $p<0.05$

Supplementary table S4 Gene sequences of TV-miR-34a plasmid.

No.	Gene sequences
1~	CGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCC
120	AGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAGCGG
121~	CCGCGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGT
240	GAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTAT
241~	AAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGG
360	AGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTA
361~	TGATCATGAACAGACTGTGAGGACTGAGGGGCCTGAAATGAGCCTTGGGACTGTGAATTTAA
480	AATACACAAACAATTAGAATCAGTAGTTTAACACATTATACACTTAAAAATTTTATAT
481~	TTACCTTAGAGCTTTAAATCTCTGTAGGTAGTTTGTCCAATTATGTCACACCACAGAAGTAAG
600	GTTCCTTCACAAAGATCCCAAGCTGTCGATCGACATTTCTAGAGGATCTCGGACCCG
601~	GGGAATCCCCGTCCCCCAACATGTCCAGATCGAAATCGTCTAGCGCGTCGGCATGCGCCATC
720	GCCACGTCCTCGCCGTCTAAGTGGAGCTCGTCCCCCAGGCTGACATCGGTCGGGGGGG
721~	CGGATCTCGGACCCGGGGAATCCCCGTCCCCCAACATGTCCAGATCGAAATCGTCTAGCGCG
840	TCGGCATGCGCCATCGCCACGTCCTCGCCGTCTAAGTGGAGCTCGTCCCCCAGGCTGA
841~	CATCGGTCGGGGGGGCGGATCCCCCGGGCTGCAGGAATTCCGGCGATACAGTCAACTGTCTT
960	TGACCTTTGTTACTACTCTCTTCCGATGATGATGTCGCACTTATTCTATGCTGTCTCA
961~	ATGTTAGAGGCATATCAGTCTCCACTGAAGCCAATCTATCTGTGACGGCATCTTTATTCACAT
1080	TATCTTGTACAAATAATCCTGTTAACAATGCTTTTATATCCTGTAAAGAATCCATTT
1081~	TCAAAATCATGTCAAGGTCTTCTCGAGGAAAAATCAGTAGAAATAGCTGTTCCAGTCTTTCTA
1200	GCCTTGATTCCACTTCTGTCAGATGTGCCCTAGTCAGCGGAGACCTTTTGGTTTTGG
1201~	GAGAGTAGCGACACTCCCAGTTGTTCTTCAGACACTTGGCGCACTTCGGTTTTTCTTTGGAGC
1320	ACTTGAGCTTTTTAAGTCGGCAAATATCGCATGCTTGTTCGATAGAAGACAGTAGCT
1321~	TCATCTTTCAGGAGGCTAGGGCCGCCCCACGTGCGCAGCAGGACGCAGCGCTGCCTGAAACT
1440	CGCGCCGCGAGGAGAGGGCGGGGCCGCGGAAAGGAAAGGGGGGGGCTGGGAGGCCCGG
1441~	AGGGGGCTGGGCCGGGGACCCGGGAGGGGTCGGGACGGGGCGGGGTCCGCGCGGAGGAGG
1560	CGGAGCTGGAAGGTGAAGGGGCAGGACGGGTGCCCGGGTCCCCAGTCCCTCCGCCACGTG
1561~	GGGAGCGCGGTCCTGGGCGTCTGTGCCCGCGAATCCACTGGGAGCCCGGCCTGGCCCCGACA
1680	GCGCAGCTGCTCCGGGCGGACCCGGGGGTCTGGGCCGCGCTTCCCCGCCCGCGCGCCG
1681~	CTCGCGCTCCCAGGGTGCAGGGACGCCAGCGAGGGCCCCAGCGGAGAGAGGTCGAATCGGC
1800	CTAGGCTGTGGGGTAACCCGAGGGAGGGGCCTCTAGATATAAGGGCGAATTCCAGCACA
1801~	CTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCCAGCTGGGAATAGAGATAGGAGGGGA
1920	CCCAGCTGGATGCAGTGGGCAGTGGGGGTCATAGAGTCAAGAGGGTACAGAATACAAT
1921~	GGGGTCCTAGTATCATGGTGGAGGTCAGAAAGAGCCCTAAAAGAGAGGGTCAAGGTAGGAG
2040	GTTAGTGAAGGTCCACCTCCACCCTCTCCAGGACAGGGACATCAGGCCACAATTAATTT
2041~	CTCTGCAGTTGGTGAGTGGTCATGGTCTCTGGAGTCCCCAGCATCCAGAGTGTCCCTGGTCTA
2160	GTGGTCCCCCCTTTCTGAGCCACAGCCACTTTCTCCATCAAATGAGGCCAGTAATAC
2161~	CCATCCCATAGTGATGCTGTGAGGATGAGATGAGCATCTGTAAGTGCTGAAGATAATCCCTG
2280	ACACATCCCAAGCATTCAGCAGTGCAAGCATACACTTACACGGCACTCCCCAGAGCCA
2281~	GGCATGTGCTGGTGCCTCATACACGTGACCACATTTGATCGTCACAATGACCCTGTGAGGGA
2400	GACTGTGCAACAGAGGACTGACCTTGCTCAAAGACCTCAGGCGTTTCCCCTCAGAGCC
2401~	TGAGAGGTCATCTCTTTTTTTTTTTTTTTTCCTTTCTTTCTTTTTCTTTTCCATTTCTTTTTCTTT
2520	GCAAGAGGTCATCTCTAATGCTTTGGAATATCCTGCCAGATTAGAGTCCCTTT
2521~	GTTCACCTGAAGGTTTGGGCCACACCAGATAGTCTAACGGTGTGATTTGTGCTGAAGGTTTTG
2640	AGCCACACTATATCAGCTAGATTTCTAGAGCGGCCGGCCGCAATAAAATATCTTTAT
2641~	TTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGTACTAACATACGCTCTCCATCA
2760	AAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAG
2761~	GTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCATTTAGGTGACACTATAGAATACAAG
2880	CTTGCATGCCTGCAGGTCCGGAGGACAGTACTCCGCTCGGAGGACAGTACTCCGCTC
2881~	GGAGGACAGTACTCCGCTCGGAGGACAGTACTCCGCTCGGAGGACAGTACTCCGACTCTAGA
3000	GGATCCCCAGTCCTATATATACTCGCTCTGCACTTGGCCCTTTTTTACACTGTGACTG
3001~	ATTGAGCTGGTGCCGTGTCGAGTGGTGTCTCGAGATCTGGCAGTGTCTTAGCTGGTTGTCTCG
3120	AGACAACCAGCTAAGACACTGCCATTTTTGCTAGCCCTCGACAATCAACCTCTGGAT
3121~	TACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGAT
3240	ACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTC
3241~	TCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAAC
3360	GTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCC
3361~	ACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCA
3480	TCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAT
3481~	TCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGG
3600	ATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTT
3601~	CCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGA
3720	GTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGGAATTCGAGCTCGGTACGGGCTC

Continued

No.	Gene sequences
3721~	GACTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGG
3840	ACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCT
3841~	ATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAT
3960	TTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTC
3961~	TACAAATGTGGTAAAATCGATAAGGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAG
4080	TCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTT
4081~	CTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACT
4200	CGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAA
4201~	TACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCA
4320	AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC
4321~	CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA
4440	AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
4441~	CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAC
4560	GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACG
4561~	AACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG
4680	TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA
4681~	GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGG
4800	ACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA
4801~	GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA
4920	TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG
4921~	ACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC
5040	TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATG
5041~	AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTC
5160	TATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGG
5161~	AGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCA
5280	GATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAA
5281~	CTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAG
5400	TTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGT
5401~	CGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCA
5520	TGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGT
5521~	TGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCAT
5640	CCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT
5641~	GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGC
5760	AGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
5761~	TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCAT
5880	CTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA
5881~	AAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATT
6000	GAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGA
6001~	AAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGT
6120	AGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTG
6121~	CCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTT
6240	TCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTAC
6241~	GGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGA
6360	TAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGT
6361~	TCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCC
6480	GATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATT
6481~	TTAACAAAATATTAACGTTTACAATTTCCCATTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTC
6600	CGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCAC
6601~	TGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAAC
6720	CAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAA
6721~	TACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCT
6840	TCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCA
6841~	CTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAA
6960	CAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
6961~	TCATACTCTTCCT
6973	

Supplementary table S5 Information on RNA-binding proteins related with genes of AKT2, GSK3β and POMC. (RBP-mRNA predicted by miRWalk 2.0)

Gene	Entrez ID	Refseq ID	RNA-binding proteins	Parclip5	Clipseq 5	iClip 5	Hits-clip5	Parclip CDS	Clipseq CDS	iClip CDS	Hits-clip CDS	Parclip3	Clipseq 3	iClip3	Hits-clip3
AKT2	208	NM_001243027	AGO2	1	0	0	0	1	1	0	1	1	1	0	1
AKT2	208	NM_001243028	AGO2	1	0	0	0	1	1	0	1	1	1	0	1
AKT2	208	NM_001626	AGO2	0	0	0	0	1	1	0	1	1	1	0	1
AKT2	208	NM_001243027	FUS	1	0	0	0	0	0	0	1	1	0	0	1
AKT2	208	NM_001243028	FUS	1	0	0	0	0	0	0	1	1	0	0	1
AKT2	208	NM_001626	FUS	0	0	0	0	1	0	0	1	1	0	0	1
AKT2	208	NM_001243028	LIN28A	0	0	0	0	1	0	0	1	1	0	0	1
AKT2	208	NM_001243027	LIN28A	0	0	0	0	1	0	0	1	1	0	0	1
AKT2	208	NM_001626	LIN28A	0	0	0	0	1	0	0	1	1	0	0	1
AKT2	208	NM_001626	FMR1	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	EIF4A3	0	0	0	1	0	0	0	1	0	0	0	1
AKT2	208	NM_001243028	EIF4A3	0	0	0	1	0	0	0	1	0	0	0	1
AKT2	208	NM_001243028	RBM10	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	RBM10	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	EIF4A3	0	0	0	1	0	0	0	1	0	0	0	1
AKT2	208	NM_001243028	HuRMNase	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	HuRMNase	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	RBM10	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	IGF2BP123	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	AGO2MNase	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	ELAVL1-MNASE	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	IGF2BP123	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	HuRMNase	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	AGO2MNase	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	ELAVL1-MNASE	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	ELAVL1-MNASE	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	FMR1	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	FMR1	1	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	CAPRIN1	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	C22ORF28	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	LIN28B	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	CAPRIN1	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	FXR2	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	ZC3H7B	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	DGCR8	0	0	0	0	0	0	0	1	0	0	0	1
AKT2	208	NM_001243027	FXR2	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	CAPRIN1	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	ZC3H7B	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	DGCR8	0	0	0	0	0	0	0	1	0	0	0	1
AKT2	208	NM_001243028	HNRNPL	0	0	0	0	0	1	0	0	0	1	0	0
AKT2	208	NM_001243027	AGO1234	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	HNRNPL	0	0	0	0	0	1	0	0	0	1	0	0
AKT2	208	NM_001626	DGCR8	0	0	0	0	0	0	0	1	0	0	0	1
AKT2	208	NM_001626	FXR2	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	ZC3H7B	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	AGO1234	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	HNRNPL	0	0	0	0	0	1	0	0	0	1	0	0
AKT2	208	NM_001243028	AGO1234	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	AGO2MNase	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001626	IGF2BP123	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	LIN28B	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	C22ORF28	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243027	LIN28B	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	C22ORF28	0	0	0	0	1	0	0	0	1	0	0	0
AKT2	208	NM_001243028	METTL3	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243028	TIAL1	0	0	0	0	0	0	0	0	0	0	1	0
AKT2	208	NM_001243027	METTL3	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243027	TIAL1	0	0	0	0	0	0	0	0	0	0	1	0
AKT2	208	NM_001626	TIA1	0	0	0	0	0	0	0	0	0	0	1	0
AKT2	208	NM_001243028	MOV10	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243027	MOV10	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001626	METTL3	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001626	TIAL1	0	0	0	0	0	0	0	0	0	0	1	0
AKT2	208	NM_001243028	PUM2	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243027	PUM2	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001626	MOV10	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243028	HuR	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243027	Ago2	0	1	0	0	0	0	0	0	0	0	0	0
AKT2	208	NM_001243027	ELAVL1	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243027	HuR	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001626	PUM2	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243028	ELAVL1	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243028	RBPMS	0	0	0	0	1	0	0	0	0	0	0	0
AKT2	208	NM_001243027	ELAVL1A	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001626	Ago2	0	1	0	0	0	0	0	0	0	0	0	0
AKT2	208	NM_001243027	RBPMS	0	0	0	0	1	0	0	0	0	0	0	0
AKT2	208	NM_001626	ELAVL1	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001626	HuR	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243028	ELAVL1A	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001243028	SFRS1	0	0	0	0	0	1	0	0	0	0	0	0

Continued

Gene	Entrez ID	Refseq ID	RNA-binding proteins	Parclip5	Clipseq 5	iClip 5	Hits-clip5	Parclip CDS	Clipseq CDS	iClip CDS	Hits-clip CDS	Parclip3	Clipseq3	iClip3	Hits-clip3
AKT2	208	NM_001243027	SFRS1	0	0	0	0	0	1	0	0	0	0	0	0
AKT2	208	NM_001626	ELAVL1A	0	0	0	0	0	0	0	0	1	0	0	0
AKT2	208	NM_001626	RBPMS	0	0	0	0	1	0	0	0	0	0	0	0
AKT2	208	NM_001243028	TDP-43	0	0	0	0	0	0	0	0	0	0	1	0
AKT2	208	NM_001243027	C17ORF85	0	0	0	0	1	0	0	0	0	0	0	0
AKT2	208	NM_001243027	EWSR1	0	0	0	1	0	0	0	0	0	0	0	0
AKT2	208	NM_001243027	TDP-43	0	0	0	0	0	0	0	0	0	0	1	0
AKT2	208	NM_001626	SFRS1	0	0	0	0	0	1	0	0	0	0	0	0
AKT2	208	NM_001243028	C17ORF85	0	0	0	0	1	0	0	0	0	0	0	0
AKT2	208	NM_001243028	EWSR1	0	0	0	1	0	0	0	0	0	0	0	0
AKT2	208	NM_001243028	TIA1	0	0	0	0	0	0	0	0	0	0	1	0
AKT2	208	NM_001626	C17ORF85	0	0	0	0	1	0	0	0	0	0	0	0
AKT2	208	NM_001243027	TIA1	0	0	0	0	0	0	0	0	0	0	1	0
AKT2	208	NM_001626	EWSR1	0	0	0	0	0	0	0	1	0	0	0	0
AKT2	208	NM_001626	TDP-43	0	0	0	0	0	0	0	0	0	0	1	0
GSK3B	2932	NM_002093	AGO2	1	0	0	1	1	1	0	0	1	1	0	1
GSK3B	2932	NM_001146156	AGO2	1	0	0	1	1	1	0	0	1	1	0	1
GSK3B	2932	NM_001146156	FUS	1	0	0	1	1	0	0	1	0	0	0	1
GSK3B	2932	NM_002093	FUS	1	0	0	1	1	0	0	1	0	0	0	1
GSK3B	2932	NM_002093	DGCR8	0	0	0	1	0	0	0	1	0	0	0	1
GSK3B	2932	NM_001146156	DGCR8	0	0	0	1	0	0	0	1	0	0	0	1
GSK3B	2932	NM_002093	EIF4A3	0	0	0	1	0	0	0	1	0	0	0	1
GSK3B	2932	NM_001146156	EIF4A3	0	0	0	1	0	0	0	1	0	0	0	1
GSK3B	2932	NM_002093	HuRMNase	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	RBM10	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	HuRMNase	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	RBM10	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	AGO2MNase	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	ELAVL1-MNASE	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	AGO2MNase	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	ELAVL1-MNASE	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	FMR1	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	LIN28A	0	0	0	0	0	0	0	1	1	0	0	1
GSK3B	2932	NM_001146156	FMR1	1	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	LIN28A	0	0	0	0	0	0	0	1	1	0	0	1
GSK3B	2932	NM_002093	AGO1234	0	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	AGO1234	0	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	ELAVL1	0	0	0	0	0	0	0	0	1	1	0	0
GSK3B	2932	NM_002093	HuR	0	0	0	0	0	0	0	0	1	1	0	0
GSK3B	2932	NM_001146156	ELAVL1	0	0	0	0	0	0	0	0	1	1	0	0
GSK3B	2932	NM_001146156	HuR	0	0	0	0	0	0	0	0	1	1	0	0
GSK3B	2932	NM_002093	IGF2BP123	0	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	RBPMS	0	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	IGF2BP123	0	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	RBPMS	0	0	0	0	1	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	C22ORF28	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_001146156	LIN28B	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	TDP-43	0	0	0	0	0	0	0	0	0	0	1	0
GSK3B	2932	NM_002093	CAPRIN1	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_002093	FXR1	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_002093	METTL3	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_002093	TIA1	0	0	0	0	0	0	0	0	0	0	1	0
GSK3B	2932	NM_001146156	CAPRIN1	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_001146156	FXR1	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_001146156	METTL3	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_001146156	TIA1	0	0	0	0	0	0	0	0	0	0	1	0
GSK3B	2932	NM_002093	FXR2	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_002093	MOV10	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	TIAL1	0	0	0	0	0	0	0	0	0	0	1	0
GSK3B	2932	NM_001146156	FXR2	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_001146156	MOV10	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	HNRNPC	0	0	0	0	0	0	0	0	0	0	1	0
GSK3B	2932	NM_002093	PUM2	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	ZC3H7B	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	HNRNPC	0	0	0	0	0	0	0	0	0	0	1	0
GSK3B	2932	NM_001146156	PUM2	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	Ago2	0	1	0	0	0	0	0	0	0	0	0	0
GSK3B	2932	NM_002093	QKI	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	Ago2	0	1	0	0	0	0	0	0	0	0	0	0
GSK3B	2932	NM_001146156	QKI	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	ELAVL1A	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	TIAL1	0	0	0	0	0	0	0	0	0	0	1	0
GSK3B	2932	NM_001146156	ELAVL1A	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_001146156	ZC3H7B	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	C17ORF85	1	0	0	0	0	0	0	0	0	0	0	0
GSK3B	2932	NM_002093	SFRS1	0	1	0	0	0	0	0	0	0	0	0	0
GSK3B	2932	NM_001146156	C17ORF85	1	0	0	0	0	0	0	0	0	0	0	0
GSK3B	2932	NM_001146156	SFRS1	0	1	0	0	0	0	0	0	0	0	0	0
GSK3B	2932	NM_002093	C22ORF28	0	0	0	0	1	0	0	0	0	0	0	0
GSK3B	2932	NM_002093	LIN28B	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	2932	NM_002093	TDP-43	0	0	0	0	0	0	0	0	0	0	1	0
POMC	5443	NM_001035256	FUS	0	0	0	1	0	0	0	0	0	0	0	0

Supplementary table S6 Distributions of C22ORF28 expressions among breast carcinoma tissues (stage I-II and stage III), their adjacent healthy breast tissue (HBT), and miR-34a status.

Different tissues	Total n (%)	C22ORF28 status		p value
		Negative (%)	Positive (%)	
miR-34a status				<0.001
Low expression	67(100)	9(13.4)	58(86.6)	
High expression	67(100)	50(74.6)	17(25.4)	
Pathological staging				<0.001
HBT	83(100)	71(85.5)	12(14.5)	
Stage I-II	74(100)	38(51.4)	36(48.6)	
Stage III	60(100)	21(35.0)	39(65.0)	

"Low expression " refers to value less than median; "High expression" refers to value more than median.

Supplementary methods

Isolation and passage of long-term-cultured BCSC

Tumor tissue-derived BCSC (XM322 and XM607) were isolated by fluorescence-activated cell sorting (FACS) as previously described [1]. Cell lines-derived BCSC (MDA-MB-231.SC and MCF-7.SC) were purified by magnetic-activated cell sorting (MACS). The dissociated cells were stained with antibodies against CD44 (Cell Signaling Inc), CD24 (Cell Signaling Inc), ALDH1 (Cell Signaling Inc) and Lin (eBioscience). Cell sorting for BCSC was performed using MACS according to the manufacturer's instructions (Miltenyi Biotec). Magnetic separation was performed up to three times to obtain a stem-like population more than 95% pure. We had described the application of culture medium previously [1]. Briefly, isolated monoclonal CSCs were maintained as spheres in ultralow attachment flasks in serum-free DMEM-F12 and supplemented with 10 ng/mL basic fibroblast growth factor (BFGF), 20 ng/mL epidermal growth factor (EGF), 2% B27, and 5 μg/mL insulin. The procedure of long-term maintenance of BCSC, briefly, included taking flask out of incubator, collecting cells, centrifuging at 2000 x *g* for 5 min, aspirating supernatant, resuspending in culture medium, aliquoting appropriate volume (1:3) of cell suspension into new flasks with media, well mixing, incubating at 37 ^{0}C, and replacing with fresh culture medium every 3-5 days.

Patients, tissues, tissue microarray construction (TMA) and immunohistochemistry (IHC)

A total of 134 female patients who hospitalized in SYSUCC from 2001 to 2006 were enrolled in our study. A complete patient follow-up was performed, and endpoint of this follow-up was January 2017. Expression data were obtained from 217

fresh-frozen resected breast specimens consisting of 134 tumor tissues and 83 adjacent healthy breast tissues (HBT). Four fresh tissues of breast tumors underwent eradicative operation in Sun Yat-Sen University Cancer Center in July 2015 were randomly chosen for the study of Western blotting. We added rabbit polyclonal primary antibody of C22ORF28 (1:100) for incubation. Staining procedures were performed by using Bench Mark XT automated IHC/ISH slide staining system.

Constructs

Instructions regarding constructions of VISA plasmid, hTERT promoter-driven VISA nanoparticle delivery of miR-34a (TV-miR-34a) and hTERT promoter-driven VISA nanoparticle delivery of control (TV-miR-Ctrl) were described previously according to the standard molecular cloning protocol [2, 3]. The hTERT promoter was amplified by PCR using genomic DNA. GAL4-VP2 contains two VP16 activation domains, amino acids 413 to 454, immediate early transactivator domain fused to the GAL4 DNA-binding domain. The mature human miR-34a sequences were obtained from the Sanger Center miRNA Registry. G5E4T contains five tandem copies of the 17-bp near-consensus DNA binding sites to GAL4 combined with the adenovirus E4 TATA box. The 800-bp WPRE fragment was released from pGEM-3Z-WPRE by Asp718/SalI digestion and incorporated into the SmaI site of the pGL3-basic vector by blunt ligation to produce intermediate pGL3-LucWPRE. In brief, the miR-34a shRNA was incorporated into the Bgl II/Nhe I sites of the plasmid pGL3-hTERT-VISA-Luc; following, the hTERT-VISA-miR-34a fragment of pGL3- hTERT -VISA-miR-34a was subcloned into the Not I and Sal I sites of pUK21. The shRNAs against green fluorescent protein (GFP) were designed and combined with hTERT-VISA in a similar manner, to create the negative control hTERT-VISA-miR-Ctrl plasmid. All

products were verified by DNA sequencing. Plasmids were amplified in DH5A *Escherichia coli* according with manufacturer's protocols. Therapeutic plasmids were purified by Qiagen Endo-Free Mega Prep Kit (Qiagen) in accordance with the manufacturer's instruction.

C22ORF28-Ad (addition of C22ORF28) and C22ORF28-KD (knockdown of C22ORF28) were generated as previously described [4, 5]. For site-specific mutagenesis, we mutated the regions in the C22ORF28-3'UTR and LIN28A-3'UTR complementary to the seed sequence of miR-34a using the QuikChange II Site-Directed Mutagenesis Kit. Mutation for C22ORF28: forward primer 5′-GAUGGGUAGAUGUCAAUGACGCUUACGCAGUCAUACUG-3′, reverse primer 5′-GUCAUACUGACGCAUUCGCAGUAACUGUAGAUGGGUAG-3′; Mutation for LIN28A: forward primer 5′-AUUGGGGCUAGUUGGCUGACGCUGUAUCUC AGGCUUGG-3′, reverse primer 5′-GGUUCGGACUCUAUGUCGCAGUCGGUUG AUCGGGGUUA-3′. Luciferase assays were carried out for 48 hr with the Dual Luciferase Reporter Assay System (Promega) according to manufacturer's protocol.

Flow cytometry

FACS was performed to analyze the $CD44^{+}CD24^{-}$ subpopulation, proportion of Lin and population of ALDH1 in BCSC by using antibodies of $CD44^{+}$, $CD24^{-}$, Lin-PE and ALDH1-PE. Regarding preparation for evaluating miR-34a expression following TV-miR-34a transduction, initially, we transfected BCSC with GFP-labeled Ctrl (control, empty vector), TV-miR-Ctrl, or TV-miR-34a plasmid for 48hr. Next, positive GFP-labeled cells were purified by FACS for further qRT-PCR analysis. Order to show the representative images, cells were contained with 2 μl DAPI (10 μg/mL) for 5 min. To determine the $CD44^{+}CD24^{-}$ population of long-term-cultured

BCSC following transfections of Ctrl, miR-Ctrl, miR-34a, TV-miR-Ctrl and TV-miR-34a, respectively, $CD44^{+}CD24^{-}$ were detected by FACS on 0, 1, 3 and 5 days following the transduction.

Serial passages of luciferase-labeled and green fluorescent protein (GFP)-labeled BCSC lineages

Establishment protocols were described previously [1, 2]. Briefly, BCSC were transfected with pEF1a-Luc-Neo, and filtrated with G418 for 14 days. Next, G418-resistant clones, referred as luciferase-labeled BCSC, were collected and maintained for further *in vivo* experiments.

Similarly, BCSC were transfected with pcDNA3.1-EGFP-NEO plasmid. Then GFP-labeled BCSC were selected out with G418 for 14 days. Following, GFP expression and survival BCSC were cultured. Here, G418-resistant clones were designated as GFP-labeled BCSC. GFP-labeled BCSC were transfected with Ctrl, TV-miR-Ctrl and TV-miR-34a, as well as TV-miR-34a co-transfected with either C22ORF28-3'UTR or C22ORF28-3'UTR-mutation (C22ORF28-3'UTR-mut). We used a spinning disk confocal long-term live cell imaging system (Olympus CV1000) to maintain and photograph these cells in real time. Mammospheres with 50 μm or greater in diameter were determined. For the detection of synergistic effects of TV-miR-34a and docetaxel, BCSC were treated with Ctrl, 1 nM docetaxel (Aventis Pharma, France), presence or absence of docetaxel after 0.1 nM TV-miR-34a transfection during in the corresponding period (0 day, 1 day, 3 days and 5 days) respectively. GFP intensity was measured with MetaMorph image acquisition and analysis software (Molecular Devices). All experiments were repeated for five times.

Clonogenicity assay in soft agarose

We performed clonogenicity assays in soft agarose to determine clonal expansion ability of BCSC as previously described [1]. Briefly, BCSC were transfected with Ctrl, TV-miR-Ctrl and TV-miR-34a; along with Ctrl, C22ORF28-Ad and C22ORF28-KD, respectively. Following, 2% solidified agarose was paved as base agar in 6-well. BCSC were seeded at 3×10^3 cells per well coated with a thin layer of 1% soft agarose. The experiment was terminated at day 21, and wells were Giemsa-stained. Spheres with 50 μm or greater in diameter were evaluated. We performed all experiments for five times.

MTT assay of cell proliferation

Cells were seeded into 96-well plates (5×10^3 cells/well), treated with Ctrl, 1 nM docetaxel (Aventis Pharma, France), presence or absence of docetaxel after 0.25 μg TV-miR-34a transfection during in the corresponding period (0 day, 1 day, 3 days and 5 days). MTT (Sigma) assay was used to assess the inhibitory effect of different interferences on the viability of various breast cells. Briefly, regarding evaluation of cell viabilities among breast cancer cells (BC, contained MDA-MB-231, MCF-7, MDA-MB-468 and SK-BR-3), BCSC (contained MDA-MB-231.SC, MCF-7.SC, XM322 and XM607) and immortal healthy mammary epithelial cells (HME, 184A1 and MCF-12A), cells (5×10^3) were transfected with 0.25 μg TV-miR-34a for 48 hr. Then, MTT was added to each well (96-well) for 4 hr. The outcomes were evaluated according to the manufacturer's instructions.

In the investigation of synergistic effects of TV-miR-34a plus docetaxel on BCSC, we maintained XM322 cells (5×10^3, 96-well) 24 hr before transduction, and we then transfected the cells with 0.25 μg TV-miR-34a and cultured with 1 nM docetaxel. Each assay was repeated at least three times.

Tumor transplantation experiment

To determine the optimum antitumor dose of T-VISA-miR-34a plasmid *in vivo*, a suspension of luciferase-labeled BCSC (1×10^4) was inoculated at the left fourth inguinal mammary gland of female BALB/c-nude mice (6-week-old; Vital River Laboratories Animal, Beijing, China). When the tumors gained ~50 mm^3, the mice were noninvasively imaged using the IVIS (*In Vivo* Imaging System, Xenogen, Alameda, CA) to confirm tumor growth and then randomly divided into four treatment groups (10 mice per group). Each group of mice received 100 μL of DNA-liposome complexes that contained 5 μg TV-miR-34a, 10 μg TV-miR-34a, 20 μg TV-miR-34a, or liposomal complexes administered through tail vein injection, every other day / quaque omni die (qod) for 4 consecutive weeks.

Moreover, to investigate the antitumor effect of TV-miR-34a *in vivo*, luciferase-labeled BCSC (1×10^4) were injected into the left fourth inguinal mammary gland of mice. When the tumors reached ~50 mm^3, the mice were noninvasively imaged using the IVIS system to confirm tumor growth and then randomly assigned to one of three following treatment groups (10 mice per group): Each group of mice received 100 μL of DNA-liposome complexes that contained TV-miR-Ctrl liposomal complexes (10 μg qod), TV-miR-34a liposomal complexes (10 μg qod) or liposomal (Ctrl) alone. The experiment was terminated on day 50. For observation of mice survival, all mice were evaluated for 80 days.

On 0, 2, 4, 6, 8 and 10 days after the injection, the mice were anesthetized and blood was collected by retro-orbital bleeding using a heparinized microcapillary tube. The concentrations of serum alanine transaminase (ALT), aspartate transaminase (AST), blood urea nitrogen (BUN) and creatinine (Cr) were determined with an

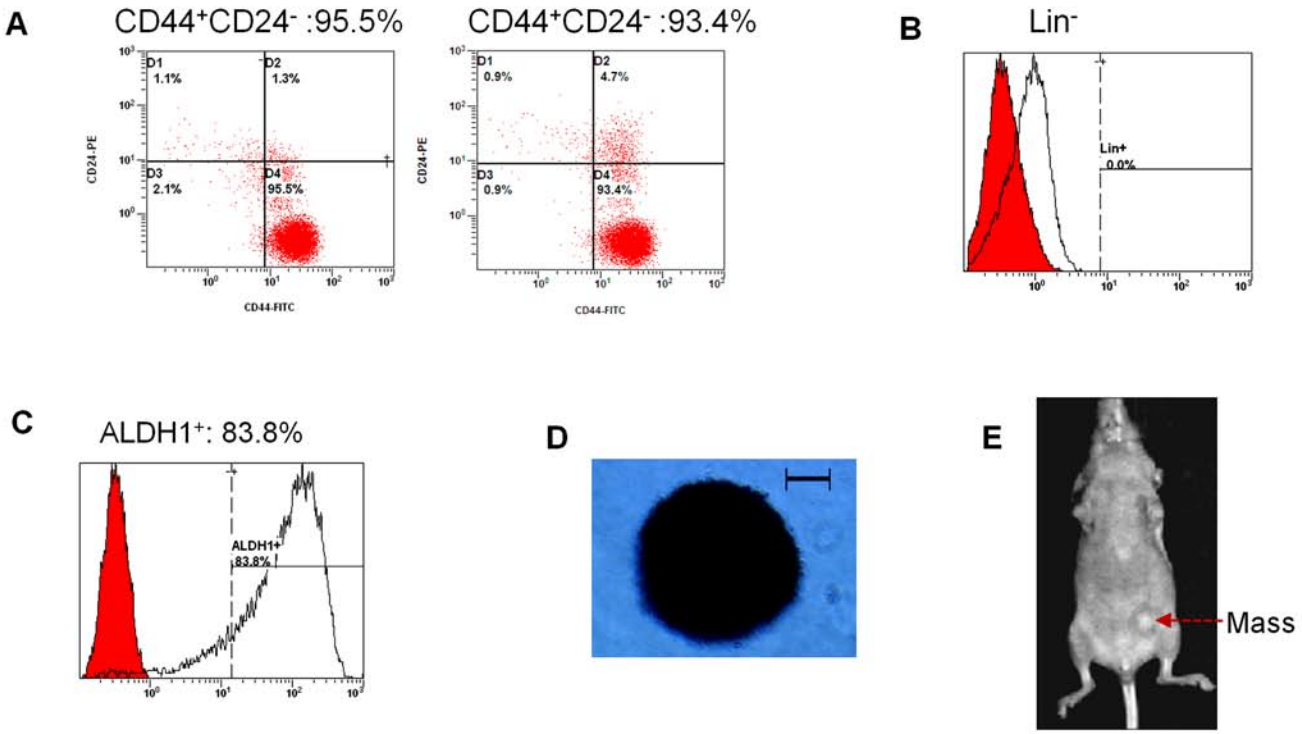
automatic analyzer (Roche Cobas Mira Plus; Roche, Mannheim, Germany). The test was repeated ten times.

The amount of cytokines (TNF-α, IL-6 and IFN-γ) in mouse sera was quantified using the cytometric bead array kit for mouse inflammatory cytokines (CBA; BD Biosciences) on a FACS Calibur cytometer equipped with Cell QuestPro and CBA software (Becton Dickinson). The test was repeated five times.

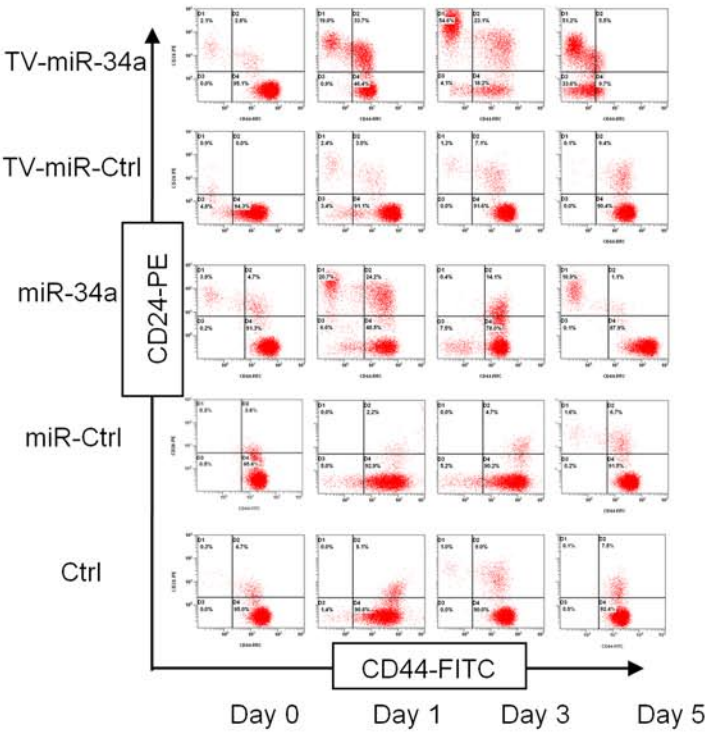
References

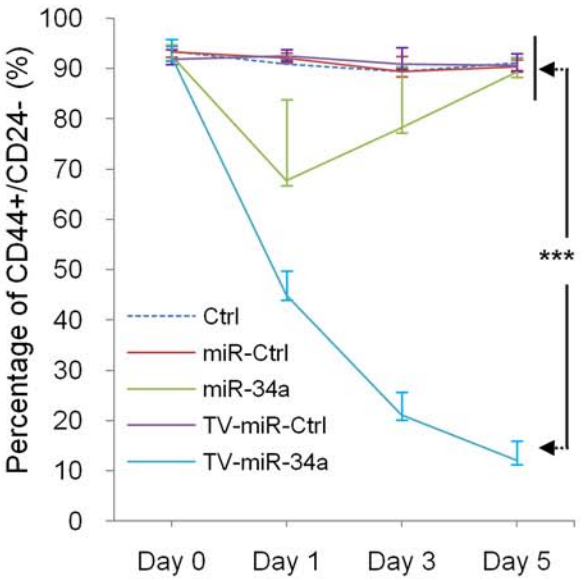
1. Lin X, Chen W, Wei F, Zhou BP, Hung MC, Xie X. POMC maintains tumor-initiating properties of tumor tissue-derived long-term-cultured breast cancer stem cells. Int J Cancer. 2017; 140: 2517-25.
2. Xie X, Xia W, Li Z, Kuo HP, Liu Y, Li Z, et al. Targeted expression of BikDD eradicates pancreatic tumors in noninvasive imaging models. Cancer Cell. 2007; 12: 52-65.
3. Li L, Xie X, Luo J, Liu M, Xi S, Guo J, et al. Targeted expression of miR-34a using the T-VISA system suppresses breast cancer cell growth and invasion. Mol Ther. 2012; 20: 2326-34.
4. Heath RJ, Leong JM, Visegrady B, Machesky LM, Xavier RJ. Bacterial and host determinants of MAL activation upon EPEC infection: the roles of Tir, ABRA, and FLRT3. PLoS Pathog. 2011; 7: e1001332-46.
5. Ray A, Zhang S, Rentas C, Caldwell KA, Caldwell GA. RTCB-1 mediates neuroprotection via XBP-1 mRNA splicing in the unfolded protein response pathway. J Neurosci. 2014; 34: 16076-85.

Supplementary figures S1-S4

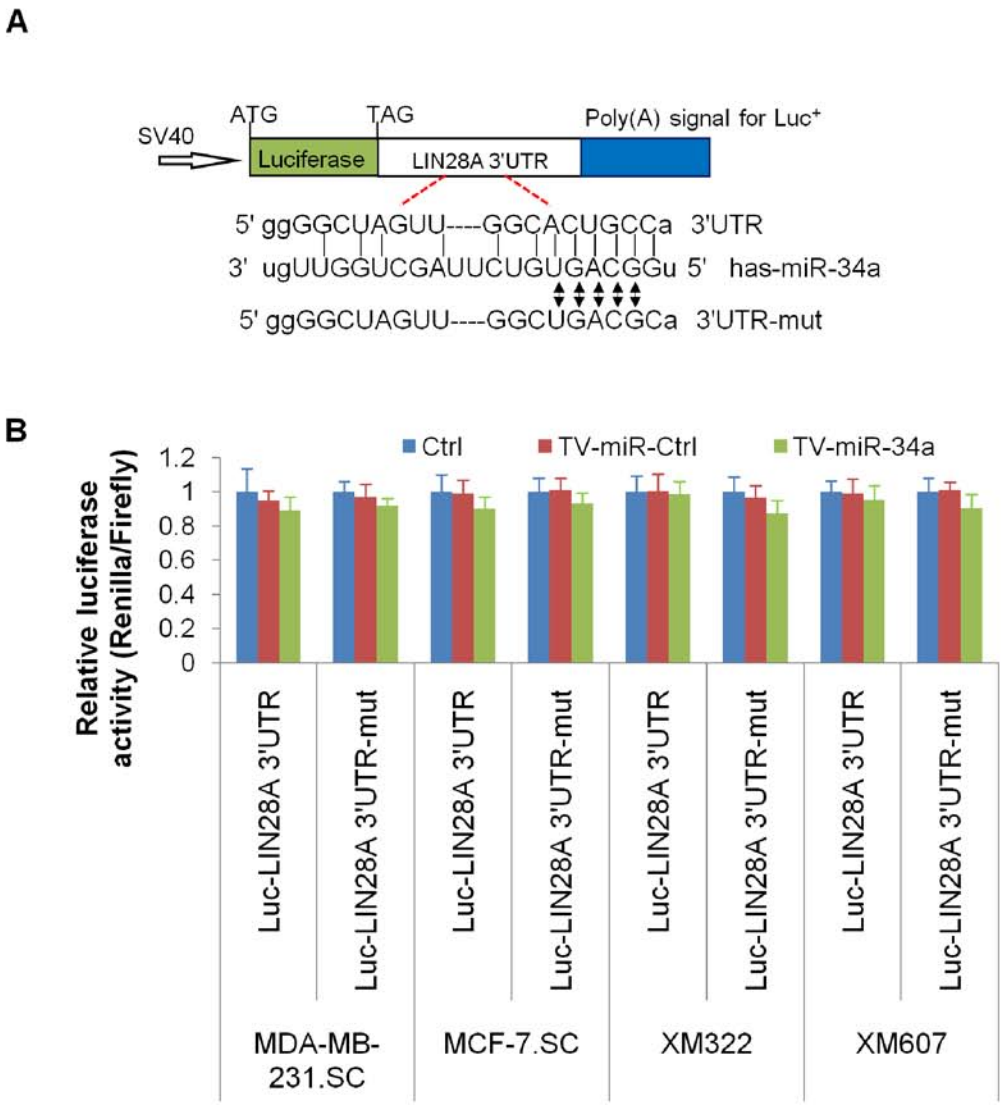


Supplementary figure S1. Tumor-initiating properties of cell line-derived BCSC can be long-term sustained. (**A**) $CD44^{+}CD24^{-}$ subpopulation in MDA-MB-231.SC (left panel) and MCF-7.SC (right panel). (**B**) $Lin^{-/low}$ expression. (**C**) $ALDH1^{+}$ marker (red represents isotype). (**D**) Clonal expansion in soft agarose. Scale bar, 100 µm. (**E**) Representative image of tumor-forming ability *in vivo* mice experiments.

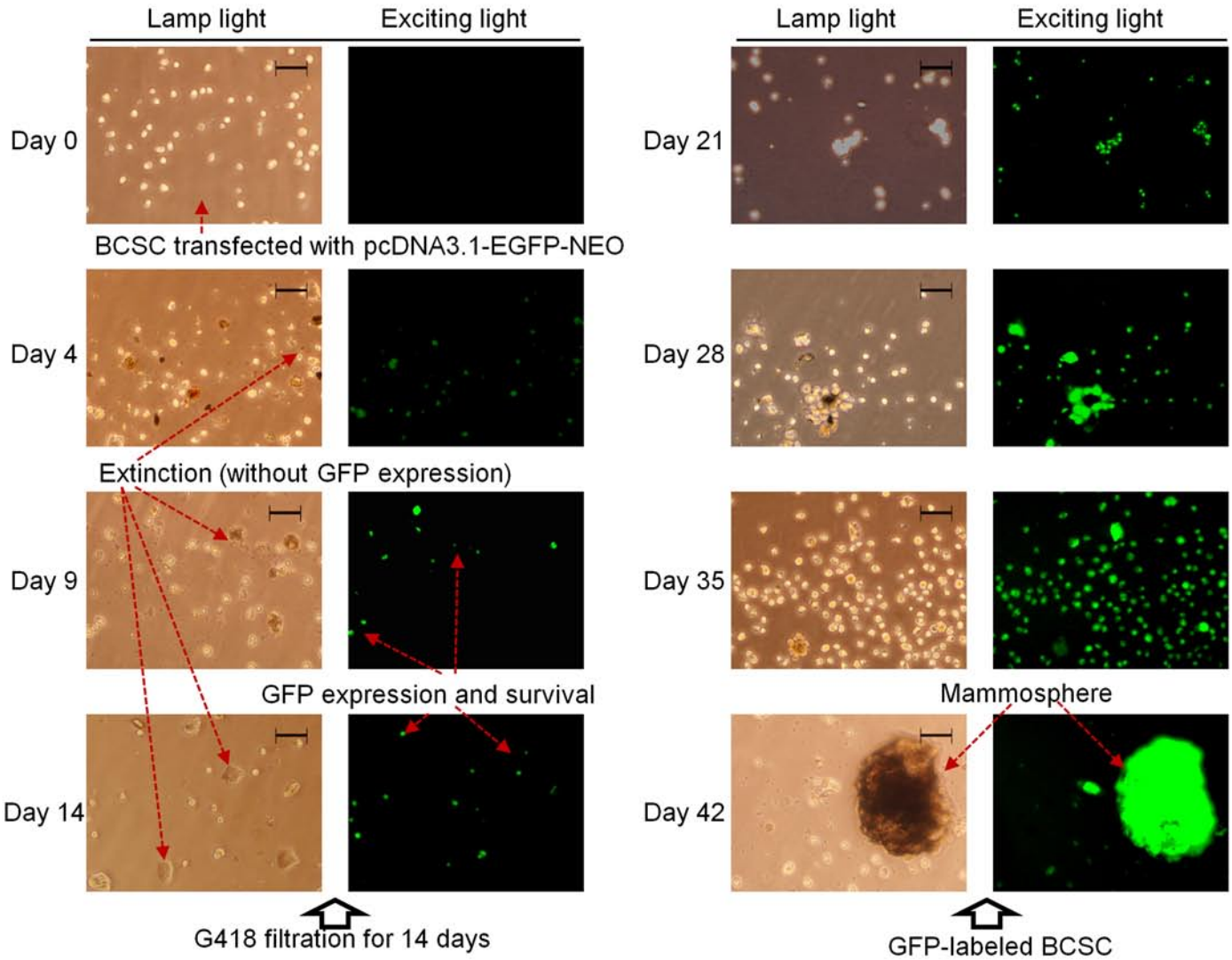




Supplementary figure S2. Representative images (left panel) and statistical results (right panel) of TV-miR-34a robustly and persistently reduced $CD44^{+}CD24^{-}$ population in MCF-7.SC; while miR-34a influence remained transient and reversible. ***$p < 0.001$.



Supplementary figure S3. Construction and determination of mutation of Luc-LIN28A-3'UTR (Luc-LIN28A-3'UTR-mut) for presence of miR-34a conversed binding sites. (**A**) Schematic diagram of Luc-LIN28A-3'UTR and Luc-LIN28A-3'UTR-mut for presence of miR-34a conversed binding sites. (**B**) Mutating the predicted miR-34a binding sites within the Luc-LIN28A-3'UTR luciferase reporter significantly abolished TV-miR-34a-dependent repression.



Supplementary figure S4. Representative images of establishment of GFP-labeled MDA-MB-231.SC. MDA-MB-231.SC initiated to express GFP protein on day 4. BCSC with GFP expression survives and proliferates; conversely, MDA-MB-231.SC without GFP expression was ruled out. Scales bar, 100 μm.