

Table S1. Detailed information on the BRCA1 germline mutation and clinicopathological characteristics of seven patients.

Patient ID	BRCA germline mutations	age(y)	parity	pathologic	ER	PR	HER2	HER2 FISH	Ki-67	PAM50 subtype	neoadjuvant therapy
Case_1	c.4013delA (p.Lys1338Argfs*28)	25	nulliparous	IDC	90%+	negative	2+	negative	20%+	Luminal A	none
Case_2	c.2194G>T (p.Glu732Ter)	43	multiparous	IMC	20%+	negative	2+	negative	70%+	Luminal B	chemotherapy
Case_3	c.1069A>T (p.Lys357Ter)	29	multiparous	IDC	negative	negative	2+	negative	80%+	Basal-like	none
Case_4	c.5251C>T (p.Arg1751Ter)	43	multiparous	IDC	negative	negative	1+	NA	60%+	Basal-like	none
Case_5	wild-type	42	multiparous	IDC	80%+	80%+	1+	NA	30%+	Luminal A	none
Case_6	wild-type	29	nulliparous	fibroadenoma	—	—	—	—	—	—	—
Case_7	wild-type	33	nulliparous	fibroadenoma	—	—	—	—	—	—	—

BRCA1 transcript: NM_007294.3

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; IDC, invasive ductal carcinoma; IMC, invasive micropapillary carcinoma; FISH, fluorescence in situ hybridization

Table S2. Cell numbers detected by single cell RNA-seq.

Patient ID	sample for scRNA-seq	histology of sample	no. of all cells	no. of normal epithelium	no. of tumor epithelium	no. of immuno	no. of fibroblast	no. of endothelium
Case_1	BC-1	breast cancer tissue	7320	205	6116	218	32	749
	ANM-1	normal glands tissue on the same side of the cancer	5734	3806	0	698	833	397
	CNM-1	normal glands tissue on the contralateral side of the cancer	7546	5278	0	1035	730	503
Case_2	BC-2	breast cancer tissue	6104	2849	487	249	1217	1302
	ANM-2	normal glands tissue on the same side of the cancer	5963	3468	99	442	906	1048
	CNM-2	normal glands tissue on the contralateral side of the cancer	5000	2457	0	386	1049	1108
Case_3	BC-3	breast cancer tissue	7845	1028	569	2859	1581	1808
	ANM-3	normal glands tissue on the same side of the cancer	7602	3018	0	1030	2665	889
Case_4	BC-4	breast cancer tissue	5938	4005	508	516	412	497
	ANM-4	normal glands tissue on the same side of the cancer	8334	3171	0	2478	1345	1340
Case_5	NC-5	normal glands tissue on the same side of the cancer	5522	4259	0	199	565	499
Case_6	NC-6	normal glands tissue on the same side of the fibroadenoma	4733	3339	0	502	304	588
Case_7	NC-7	normal glands tissue on the same side of the fibroadenoma	4481	2104	0	837	403	1137

Abbreviations: BC, breast cancer; ANM, adjacent normal mammary; CNM, contralateral normal mammary; NC, normal control.

Table S3. List of cell type specific marker genes identified at a single cell resolution**a. endothelial cell upregulated**

Gene symbol	Log2 Fold Change	P-Value	AUC
ACKR1	7.39	9.86E-231	0.83
CCL14	7.15	1.11E-216	0.83
CLDN5	7.10	4.17E-214	0.86
CXorf36	7.09	1.15E-206	0.81
PLVAP	6.94	3.43E-207	0.91
VWF	6.88	9.56E-203	0.87
ADGRL4	6.84	1.01E-202	0.92
EMCN	6.70	3.59E-195	0.91
SELE	6.47	7.38E-179	0.78
PECAM1	6.36	4.54E-178	0.95
TSPAN7	6.29	8.28E-173	0.83
RAMP3	6.17	4.07E-167	0.83
FLT1	6.05	8.17E-158	0.77
S1PR1	6.04	3.15E-160	0.82
CLEC14A	5.99	2.12E-157	0.81
SDPR	5.94	2.59E-155	0.87
AQP1	5.84	2.06E-151	0.90
C2CD4B	5.82	2.87E-146	0.73
CD93	5.77	3.38E-147	0.84
ZNF385D	5.49	8.24E-133	0.76
MCTP1	5.27	3.31E-122	0.80
RBP7	5.26	3.75E-120	0.80
RAMP2	5.24	2.30E-122	0.91
CYYR1	5.24	1.08E-121	0.82
CD34	5.09	9.88E-116	0.88
FABP4	5.03	1.59E-110	0.79
CSF3	5.03	2.87E-104	0.63
PALMD	4.96	3.62E-109	0.81
GNG11	4.95	1.12E-109	0.94
IL33	4.85	9.90E-104	0.78
EGFL7	4.69	2.91E-97	0.84
CD36	4.64	9.93E-95	0.83
THBD	4.63	6.13E-94	0.78
RNASE1	4.61	3.72E-93	0.91
ENG	4.57	3.28E-92	0.87
CALCRL	4.49	3.82E-88	0.79
ESAM	4.42	3.31E-86	0.84
ICAM2	4.39	4.82E-84	0.76
JAM2	4.34	1.45E-82	0.81
ADAMTS9	4.18	5.89E-75	0.73
NRN1	4.17	7.56E-76	0.82
IFI44L	4.01	2.07E-69	0.76
SPRY1	3.98	2.10E-68	0.79
LIFR	3.95	6.88E-67	0.77
IFI27	3.93	3.01E-67	0.93
LDB2	3.83	6.73E-63	0.80
TGFBR2	3.75	3.62E-60	0.83
KCTD12	3.71	1.91E-58	0.76
HSPG2	3.71	2.50E-58	0.76
SLC9A3R2	3.67	3.82E-56	0.65
APOLD1	3.56	5.50E-53	0.74
FABP5	3.34	1.95E-46	0.78
CDC42EP3	3.30	2.09E-45	0.78
UPP1	3.29	8.26E-45	0.71
SPARCL1	3.27	9.62E-45	0.91
LMCD1	3.23	2.68E-43	0.83
A2M	3.21	6.35E-43	0.86
PNP	3.19	3.15E-42	0.76
GJA1	3.16	4.51E-41	0.74
IL6	3.15	3.60E-40	0.69
VAMP5	3.13	1.84E-40	0.79
ICAM1	3.13	2.14E-40	0.80
CAV1	3.12	2.86E-40	0.89
CLEC3B	3.09	2.92E-38	0.71

b.fibroblast cell upregulated

Gene symbol	Log2 Fold Change	P-Value	AUC
SFRP4	7.12	1.01E-156	0.60
SFRP2	6.96	8.28E-198	0.87
DPT	6.89	2.01E-193	0.73
PI16	6.87	1.64E-166	0.58
TAC1	6.82	8.19E-181	0.63
DCN	6.63	6.24E-186	0.98
COMP	6.51	1.74E-146	0.55
IGF1	6.50	1.83E-177	0.82
LUM	6.29	5.18E-168	0.96
PTGDS	5.85	5.80E-141	0.75
CCDC80	5.82	4.93E-142	0.84
THBS2	5.78	2.22E-138	0.69
FBLN1	5.77	2.57E-140	0.82
COL1A1	5.64	1.14E-132	0.84
PDGFRL	5.56	6.02E-130	0.75
CFD	5.50	9.25E-125	0.74
SRPX	5.48	8.97E-128	0.87
COL1A2	5.47	1.04E-125	0.89
TNFAIP6	5.38	7.62E-121	0.80
COL3A1	5.34	3.99E-120	0.82
MFAP4	5.29	4.74E-118	0.83
MMP2	5.24	2.14E-115	0.84
IGFBP6	5.20	2.92E-111	0.76
CTHRC1	5.18	8.09E-110	0.62
CTSK	5.11	8.09E-110	0.90
MEG3	5.06	4.22E-107	0.80
MXRA5	5.06	4.61E-107	0.76
OLFML3	5.03	1.23E-106	0.81
CXCL14	4.94	5.21E-97	0.68
C1S	4.92	2.57E-102	0.91
APOD	4.88	2.46E-99	0.89
COL6A3	4.81	8.99E-97	0.86
LRP1	4.75	9.54E-95	0.86
ISLR	4.70	3.13E-92	0.80
COL5A2	4.62	3.28E-88	0.75
COL5A1	4.61	5.44E-87	0.68
SPON2	4.54	3.57E-85	0.81
SERPINF1	4.53	2.88E-85	0.85
VCAN	4.52	8.75E-85	0.81
EFEMP1	4.46	8.39E-82	0.77
POSTN	4.43	1.65E-79	0.66
SERPINE2	4.41	4.75E-79	0.73
FBLN5	4.37	6.79E-78	0.75
C1R	4.34	2.85E-77	0.76
RARRES2	4.29	9.73E-76	0.85
AEBP1	4.29	5.31E-75	0.77
FBN1	4.28	5.26E-74	0.69
CD248	4.27	6.69E-74	0.72
COL12A1	4.21	1.91E-70	0.66
HTRA1	4.18	4.61E-71	0.74
C3	4.14	8.54E-68	0.67
MXRA8	4.10	5.34E-68	0.75
PCOLCE	4.07	2.19E-67	0.84
MMP11	4.02	4.50E-55	0.54
PLXDC1	3.99	2.34E-63	0.67
COL6A1	3.99	1.96E-64	0.84
COL6A2	3.91	9.79E-62	0.88
COL14A1	3.90	1.52E-60	0.70
FN1	3.90	2.49E-59	0.75
MMP19	3.83	4.81E-57	0.70
ASPN	3.82	3.29E-55	0.59
CFH	3.80	4.81E-57	0.75
CXCL12	3.78	7.55E-57	0.87
FSTL1	3.74	9.97E-56	0.80
EFEMP2	3.73	3.95E-55	0.73
NID1	3.66	1.31E-52	0.74
ELN	3.65	1.61E-45	0.58
GEM	3.60	4.17E-51	0.84
CTGF	3.58	4.66E-50	0.80
PLTP	3.54	1.23E-48	0.74
THBS4	3.50	8.15E-41	0.53

Table S3. List of cell type specific marker genes identified at a single cell resolution**c. luminal progenitor cell upregulated**

Gene symbol	Log2 Fold Change	P-Value	AUC
PI3	6.44	2.84E-151	0.64
PIGR	5.70	1.91E-143	0.83
GABRP	5.22	4.21E-122	0.87
C2orf82	5.16	1.76E-117	0.73
ALDH1A3	4.69	1.53E-99	0.89
MMP7	4.66	4.35E-95	0.81
LIF	4.50	7.02E-91	0.86
CCL28	4.44	1.03E-88	0.88
ITF	4.15	2.75E-73	0.83
KRT23	4.14	1.95E-76	0.87
KRT6B	4.13	4.50E-75	0.84
TPT1-AS1	3.83	1.51E-65	0.91
SAA1	3.73	1.58E-59	0.79
S100A2	3.37	1.22E-42	0.62
PDE4B	3.13	3.33E-42	0.89
CXCL3	3.07	1.56E-40	0.88
RCAN1	3.06	2.26E-40	0.89
SFRP1	3.04	1.46E-39	0.88
NDRG2	3.02	3.22E-39	0.91
PTN	2.97	1.16E-37	0.67
SLPI	2.85	1.52E-33	0.90
CXCL2	2.65	1.26E-29	0.92
KRT7	2.55	2.71E-27	0.94
LAMB3	2.47	3.52E-25	0.67
SDC4	2.41	3.55E-24	0.90
HMGA1	2.35	4.97E-23	0.87
CRABP1	2.33	4.07E-22	0.78
MRPS24	2.30	4.77E-22	0.84
ITGB6	2.30	1.06E-21	0.81
S100A1	2.29	2.43E-21	0.83
SLC25A37	2.27	2.83E-21	0.84
MAP2K3	2.23	1.18E-20	0.89
NTRK2	2.23	2.19E-20	0.77
CDC42EP1	2.17	1.97E-19	0.88
TACSTD2	2.15	4.16E-19	0.92
CXCL16	2.14	9.72E-19	0.84
CXCL1	2.10	7.95E-18	0.75
MFGE8	2.04	4.43E-17	0.83
FOSL1	2.02	1.38E-16	0.81
WFDC2	2.00	2.75E-16	0.78

d. mature luminal cell upregulated

Gene symbol	Log2 Fold Change	P-Value	AUC
PIP	6.37	1.21E-163	0.72
CITED1	6.15	6.53E-162	0.64
MUCL1	6.03	7.99E-152	0.70
MYBPC1	5.97	3.40E-161	0.70
TFF1	5.95	8.53E-158	0.73
ANKRD30A	5.56	8.03E-143	0.89
PTHLH	5.11	4.39E-126	0.74
AREG	4.87	5.94E-117	0.89
SERPINA1	4.70	5.46E-100	0.65
STC2	4.65	7.67E-108	0.89
SCGB1D2	4.50	4.87E-49	0.51
TFF3	4.47	5.13E-99	0.83
AGR2	4.45	1.11E-98	0.90
EFHD1	3.83	2.24E-75	0.89
AFF3	3.82	1.61E-74	0.85
AZGP1	3.76	8.01E-73	0.95
BAMBI	3.65	1.10E-67	0.76
AGR3	3.39	5.55E-59	0.82
SCGB3A1	3.24	3.75E-41	0.65
TBX3	3.15	3.46E-51	0.78
GDF15	2.98	1.21E-42	0.61
EREG	2.98	2.91E-44	0.67
C8orf4	2.82	1.80E-41	0.78
BATF	2.80	2.94E-40	0.77
S100A14	2.79	1.37E-40	0.86
TFPI2	2.60	1.20E-32	0.53
RAB11FIP1	2.57	1.55E-34	0.80
LURAP1L	2.55	6.22E-34	0.74
UGDH	2.54	1.56E-33	0.73
MGP	2.49	1.99E-32	0.87
ELOVL5	2.48	2.33E-32	0.84
TFAP2A	2.41	3.00E-30	0.76
PLAT	2.38	1.33E-29	0.81
CXCL13	2.38	4.33E-26	0.56
KIAA1324	2.34	1.61E-28	0.82
GSTM3	2.29	1.97E-27	0.74
CRACR2B	2.28	4.46E-27	0.81
SCGB2A2	2.23	5.32E-21	0.55
RUNX1	2.21	1.57E-25	0.79
DUSP4	2.09	9.19E-23	0.69
SMIM14	2.08	1.17E-22	0.79
SPINT2	2.08	1.33E-22	0.88
SPDEF	2.06	3.99E-22	0.75

Table S3. List of cell type specific marker genes identified at a single cell resolution**e. basal progenitor cell upregulated**

Gene symbol	Log2 Fold Change	P-Value	AUC
WIF1	7.07	4.63E-132	0.68
KRT14	5.51	4.77E-72	0.91
KRT17	5.47	1.19E-70	0.91
C2orf40	5.15	4.16E-60	0.90
TNS4	4.62	3.32E-43	0.75
DST	4.45	1.01E-39	0.90
CNN1	4.41	1.78E-38	0.81
ACTG2	4.33	2.40E-35	0.71
MYLK	4.32	1.52E-36	0.82
KRT5	3.93	7.11E-28	0.89
LAMA3	3.85	6.44E-26	0.69
SFN	3.80	2.72E-25	0.86
MATN2	3.68	5.06E-23	0.71
SEMA3C	3.61	6.95E-22	0.75
DKK3	3.34	1.08E-17	0.73
ACTA2	3.23	3.26E-16	0.87
LAMB3	3.02	2.67E-13	0.73
STK17A	2.95	1.38E-12	0.65
MT1X	2.87	7.67E-12	0.76
TUBB2A	2.85	1.32E-11	0.78
SAA1	2.81	1.03E-10	0.76
SCPEP1	2.80	4.60E-11	0.73
CRYAB	2.75	1.60E-10	0.80
TPM2	2.72	2.93E-10	0.84
FBXO32	2.67	9.83E-10	0.68
GLIPR1	2.65	1.56E-09	0.66
H2AFX	2.60	3.61E-09	0.70
TAGLN	2.60	3.83E-09	0.82
PRNP	2.52	1.96E-08	0.77
TSHZ2	2.49	3.76E-08	0.67
NTRK2	2.48	4.20E-08	0.67
CSRP1	2.38	2.24E-07	0.77
ANXA1	2.38	2.40E-07	0.82
MLF1	2.29	1.02E-06	0.68
MT1E	2.29	1.05E-06	0.64
CYR61	2.28	1.12E-06	0.76
SLC1A5	2.23	2.81E-06	0.64
SGK1	2.20	4.17E-06	0.68
LBH	2.20	4.34E-06	0.64
NEDD4L	2.18	5.69E-06	0.66
PDLIM4	2.18	5.62E-06	0.68
PIK3R1	2.15	8.21E-06	0.60
CXCL14	2.14	1.90E-05	0.69
SFRP1	2.10	1.76E-05	0.68
MYH11	2.09	2.53E-05	0.64
NRP2	2.03	4.70E-05	0.58
MT2A	2.02	4.51E-05	0.68
THBS1	2.01	5.88E-05	0.68

f. myoepithelium cell upregulated

Gene symbol	Log2 Fold Change	P-Value	AUC
RERGL	8.11	5.33E-213	0.70
PLN	7.90	3.77E-206	0.73
MYH11	5.97	9.33E-116	0.75
COX4I2	5.93	4.69E-113	0.69
CCL19	5.81	7.13E-73	0.55
NDUFA4L2	5.16	4.68E-82	0.81
PPP1R14A	5.00	5.39E-76	0.79
CCL8	4.80	1.25E-62	0.62
RGS5	4.80	2.60E-67	0.78
GJA4	4.69	2.17E-63	0.71
CH25H	4.68	7.39E-59	0.59
SYNPO2	4.66	5.77E-63	0.73
MYL9	4.62	5.77E-63	0.94
MT1A	4.45	2.43E-55	0.72
ACTA2	4.33	5.62E-53	0.89
TAGLN	4.27	2.93E-51	0.92
SORBS2	4.19	3.30E-48	0.75
EDNRA	4.18	1.16E-47	0.72
ENPEP	4.04	4.54E-43	0.62
MT1M	4.04	7.18E-44	0.80
GPX3	3.98	3.15E-42	0.77
MCAM	3.94	4.74E-41	0.75
TPM2	3.93	1.88E-41	0.89
HSPB6	3.88	2.14E-39	0.72
IGFBP5	3.84	1.40E-38	0.86
NOTCH3	3.84	1.93E-38	0.77
LBH	3.78	4.15E-37	0.75
4-Sep	3.71	4.59E-35	0.69
RASL11A	3.69	2.08E-34	0.70
SOD3	3.58	3.12E-32	0.76
FHL1	3.56	1.38E-31	0.72
ACTG2	3.53	9.36E-30	0.63
CCL21	3.43	7.25E-20	0.56
COL18A1	3.35	5.36E-27	0.67
PLAC9	3.33	8.76E-27	0.81
TINAGL1	3.32	1.28E-26	0.76
CRISPLD2	3.29	8.01E-26	0.72
CNN1	3.26	5.50E-25	0.66
MAP1B	3.21	1.79E-24	0.78
FXYP1	3.17	1.48E-23	0.75
PTP4A3	3.14	8.23E-23	0.68
C11orf96	3.12	1.89E-22	0.70
MYLK	3.11	1.93E-22	0.77
CALD1	3.10	1.92E-22	0.87
ADAMTS1	3.07	1.08E-21	0.75
PRRX1	3.02	9.49E-21	0.71
ITGA1	3.02	1.47E-20	0.61
IGFBP7	3.01	8.88E-21	0.88
ANGPTL4	3.01	2.91E-20	0.63

Table S3. List of cell type specific marker genes identified at a single cell resolution

g. T cell upregulated

Gene symbol	Log2 Fold Change	P-Value	AUC	Gene symbol	Log2 Fold Change	P-Value	AUC
XCL2	8.93	2.79E-148	0.59	RAC2	5.21	8.12E-49	0.75
GZMA	8.91	5.70E-193	0.69	PARP8	5.10	3.16E-45	0.69
GZMH	8.90	3.64E-177	0.62	CD53	5.03	3.16E-43	0.67
NKG7	8.74	2.50E-195	0.71	LIMD2	5.00	1.01E-42	0.67
KLRB1	8.68	1.32E-184	0.68	CD48	4.99	2.76E-42	0.68
CD3D	8.66	5.49E-192	0.81	GNG2	4.97	3.00E-42	0.68
CD3E	8.65	1.94E-202	0.91	CD37	4.90	3.08E-40	0.78
CD8A	8.57	2.52E-193	0.73	FYN	4.86	1.80E-39	0.71
CD2	8.51	1.85E-183	0.79	PLAC8	4.81	1.35E-33	0.55
CD96	8.48	7.88E-173	0.70	ALOX5AP	4.80	1.34E-37	0.69
GZMK	8.48	4.78E-173	0.66	FCMR	4.66	4.08E-33	0.58
ICOS	8.45	2.11E-162	0.64	LCP1	4.54	4.09E-32	0.73
XCL1	8.39	9.14E-156	0.64	LEPROTL1	4.40	1.56E-29	0.77
GNLY	8.32	1.39E-161	0.60	TAGAP	4.23	2.83E-25	0.62
GZMM	8.29	2.02E-161	0.67	GPSM3	4.14	2.43E-24	0.65
CD8B	7.93	8.41E-144	0.64	WIPF1	4.11	5.04E-24	0.67
CD247	7.91	2.40E-144	0.66	SMAP2	4.08	2.43E-23	0.63
LCK	7.90	1.08E-144	0.67	ADGRE5	4.05	8.77E-23	0.64
TRBC2	7.62	9.68E-136	0.70	IRF4	4.04	1.03E-19	0.54
CD7	7.25	1.21E-127	0.83	LAPTM5	4.00	3.54E-22	0.79
CST7	7.14	5.62E-122	0.83	CREM	3.89	1.25E-20	0.73
SH2D2A	7.05	1.98E-110	0.67	HCLS1	3.89	2.62E-20	0.64
PTPRCAP	6.99	3.50E-113	0.69	IL32	3.88	2.02E-20	0.83
IL7R	6.90	3.52E-111	0.82	CD70	3.87	1.69E-15	0.51
SYTL3	6.53	5.26E-94	0.72	RGS1	3.82	4.81E-19	0.72
PTPRC	6.50	5.01E-95	0.89	STK4	3.82	2.45E-19	0.67
SPOCK2	6.50	6.34E-93	0.73	ITGB2	3.78	1.84E-18	0.64
CD69	6.49	1.44E-92	0.82	CDC42SE2	3.75	3.50E-18	0.62
GPR65	6.19	4.48E-79	0.65	TUBA4A	3.68	2.65E-17	0.64
HCST	6.19	1.46E-81	0.78	ARHGDI1B	3.65	3.09E-17	0.81
STAT4	6.18	3.76E-79	0.66	PTGER4	3.64	9.20E-17	0.62
LTB	6.04	1.41E-72	0.64	STK17B	3.60	2.47E-16	0.65
IL2RG	6.03	9.35E-76	0.79	SRGN	3.56	6.61E-16	0.88
CD52	6.03	8.29E-76	0.85	6-Sep	3.48	7.37E-15	0.63
CCL5	5.98	6.55E-74	0.86	IL10RA	3.45	5.24E-14	0.59
CXCR4	5.97	4.92E-74	0.88	TGFB1	3.38	1.29E-13	0.62
CYTIP	5.83	2.35E-68	0.78	DUSP2	3.35	2.70E-13	0.73
CLEC2D	5.70	1.74E-62	0.65	ARL4C	3.33	5.17E-13	0.63
CORO1A	5.69	1.74E-63	0.78	CELF2	3.23	5.31E-12	0.60
GZMB	5.54	3.69E-50	0.59	LY9	3.16	1.34E-09	0.53
FYB	5.28	1.22E-50	0.75	GMFG	3.14	4.09E-11	0.63
SELL	5.25	1.07E-45	0.56	RHOH	3.11	9.21E-11	0.62
SAMSN1	5.25	3.13E-49	0.69	TSC22D3	3.09	1.07E-10	0.71

Table S3. List of cell type specific marker genes identified at a single cell resolution

h. Macrophage cell upregulated

Gene symbol	Log2 Fold Change	P-Value	AUC
XCL2	8.93	2.79E-148	0.59
GZMA	8.91	5.70E-193	0.69
GZMH	8.90	3.64E-177	0.62
NKG7	8.74	2.50E-195	0.71
KLRB1	8.68	1.32E-184	0.68
CD3D	8.66	5.49E-192	0.81
CD3E	8.65	1.94E-202	0.91
CD8A	8.57	2.52E-193	0.73
CD2	8.51	1.85E-183	0.79
CD96	8.48	7.88E-173	0.70
GZMK	8.48	4.78E-173	0.66
ICOS	8.45	2.11E-162	0.64
XCL1	8.39	9.14E-156	0.64
GNLY	8.32	1.39E-161	0.60
GZMM	8.29	2.02E-161	0.67
CD8B	7.93	8.41E-144	0.64
CD247	7.91	2.40E-144	0.66
LCK	7.90	1.08E-144	0.67
TRBC2	7.62	9.68E-136	0.70
CD7	7.25	1.21E-127	0.83
CST7	7.14	5.62E-122	0.83
SH2D2A	7.05	1.98E-110	0.67
PTPRCAP	6.99	3.50E-113	0.69
IL7R	6.90	3.52E-111	0.82
SYTL3	6.53	5.26E-94	0.72
PTPRC	6.50	5.01E-95	0.89
SPOCK2	6.50	6.34E-93	0.73
CD69	6.49	1.44E-92	0.82
GPR65	6.19	4.48E-79	0.65
HCST	6.19	1.46E-81	0.78
STAT4	6.18	3.76E-79	0.66
LTB	6.04	1.41E-72	0.64
IL2RG	6.03	9.35E-76	0.79
CD52	6.03	8.29E-76	0.85
CCL5	5.98	6.55E-74	0.86
CXCR4	5.97	4.92E-74	0.88
CYTIP	5.83	2.35E-68	0.78
CLEC2D	5.70	1.74E-62	0.65
CORO1A	5.69	1.74E-63	0.78
GZMB	5.54	3.69E-50	0.59
FYB	5.28	1.22E-50	0.75
SELL	5.25	1.07E-45	0.56
SAMSN1	5.25	3.13E-49	0.69

i. B cell upregulated

Gene symbol	Log2 Fold Change	P-Value	AUC
RAC2	5.21	8.12E-49	0.75
PARP8	5.10	3.16E-45	0.69
CD53	5.03	3.16E-43	0.67
LIMD2	5.00	1.01E-42	0.67
CD48	4.99	2.76E-42	0.68
GNG2	4.97	3.00E-42	0.68
CD37	4.90	3.08E-40	0.78
FYN	4.86	1.80E-39	0.71
PLAC8	4.81	1.35E-33	0.55
ALOX5AP	4.80	1.34E-37	0.69
FCMR	4.66	4.08E-33	0.58
LCP1	4.54	4.09E-32	0.73
LEPROTL1	4.40	1.56E-29	0.77
TAGAP	4.23	2.83E-25	0.62
GPSM3	4.14	2.43E-24	0.65
WIPF1	4.11	5.04E-24	0.67
SMAP2	4.08	2.43E-23	0.63
ADGRE5	4.05	8.77E-23	0.64
IRF4	4.04	1.03E-19	0.54
LAPTM5	4.00	3.54E-22	0.79
CREM	3.89	1.25E-20	0.73
HCLS1	3.89	2.62E-20	0.64
IL32	3.88	2.02E-20	0.83
CD70	3.87	1.69E-15	0.51
RGS1	3.82	4.81E-19	0.72
STK4	3.82	2.45E-19	0.67
ITGB2	3.78	1.84E-18	0.64
CDC42SE2	3.75	3.50E-18	0.62
TUBA4A	3.68	2.65E-17	0.64
ARHGDIB	3.65	3.09E-17	0.81
PTGER4	3.64	9.20E-17	0.62
STK17B	3.60	2.47E-16	0.65
SRGN	3.56	6.61E-16	0.88
6-Sep	3.48	7.37E-15	0.63
IL10RA	3.45	5.24E-14	0.59
TGFB1	3.38	1.29E-13	0.62
DUSP2	3.35	2.70E-13	0.73
ARL4C	3.33	5.17E-13	0.63
CELF2	3.23	5.31E-12	0.60
LY9	3.16	1.34E-09	0.53
GMFG	3.14	4.09E-11	0.63
RHOH	3.11	9.21E-11	0.62
TSC22D3	3.09	1.07E-10	0.71

Table S4. EMT signature gene set

Gene set name	epithelial gene set		mesenchymal gene set	
genes	SCGB2A2	RPL6	TPM2	TGFBI
	TNF	PELI1	BASP1	IGFBP3
	CXCL2	ISG20	BGN	CDH2
	CXCL1	CXCL16	ID3	FBLN2
	CXCL3	LRRFIP2	CTHRC1	LAMC1
	LIF	RPL14	SPARC	RGS4
	SCGB1D2	EIF3L	MYLK	RARRES2
	PRSS8	KRT18	MFAP5	NTM
	NDRG2	RPS14	PRRX1	COL3A1
	NTRK2	RPS23	ELN	MMP1
	PTN	NACA	FN1	VCAM1
	LTF	RPL11	LOXL2	NID2
	MFGE8	DDIT4	ADAMTS1	LRP1
	GLRX	RPL29	PTX3	COL5A3
	GABRP	PPA1	GEM	IGFBP4
	WFDC2	RPL32	SFRP2	EDIL3
	IMPDH2	KRT23	COL4A2	FOS
	TNFAIP2	RPL7A	CXCL14	LUM
	PHLDA1	RPLP0	VEGFA	SLIT3
	MAP2K3	RPL19	FBLN1	ITGB3
	CXCL8	CSRP1	IL6	SPOCK1
	NFKBIZ	APP	COL11A1	SDC1
	PDCD4	EIF3D	MMP2	ECM2
	SDC4	RPSA	ITGB1	FAP
	ERRFI1	LMNA	GAS1	SERPINE2
	FOSL1	DNAJB6	TIMP3	LRRC15
	ZC3H12A	RPS3A	POSTN	LOX
	CD24	RPL4	COL6A2	PCOLCE
	ANXA1	C1QBP	IGFBP2	ZEB1
	TRAF4	ENO1	COL5A1	FERMT2
	SOD2	RPL12	FOXC1	CDH11
	TNFRSF12A	RPS19	SNAI2	SERPINF1
	BIRC3	RGS2	COL12A1	APLP1
	HMGA1	RPS15A	COL1A1	TGFBI
	TPT1-AS1	PRELID1	COL1A2	
	PLAU	CCT2	PLOD2	
	CLDN3	EEF1B2	PMEPA1	
	RPL3	KRT7	LOXL1	
	RCAN1	GNB2L1	SFN	
	GLIPR1	FBL	IGFBP7	
	TACSTD2	SLC3A2	GPC1	
	RPS9	CRYAB	MATN2	
	RPS24	TM4SF1	THY1	
	APEX1	FBXO32	VCAN	
	CLDN7	RPL5	COLGALT1	
	GLTSCR2	EEF1A1	MXRA5	
	EEF2	CRABP2	AEBP1	
	CLDN4	EPCAM	COL5A2	
	ATP1B1	RPS27A	LAMA2	
	RSL1D1	RPS18	WIPF1	
	RPL10A	PABPC1	ECM1	
	RPL13		GREM1	

Table S5. Clinicopathologic information of 8 additional BRCA1 mutation carriers and 14 non-carriers for bulk RNA-seq.

Sample ID	Age	BRCA1 germline mutation	Histologic	ER statuses	PR statuses	HER2 statuses	Molecular subtype
IH1NMT	34	c.3288_3289del (p.Leu1098Serfs)	IDC	Positive	Positive	Negative	Luminal
1643NMT	35	c.2899del (p.Thr967Leufs)	DCIS	Negative	Negative	Negative	TNBC
1769NMT	26	c.4013del (p.Lys1338Argfs)	IDC	Positive	Positive	Negative	Luminal
IH8NMT	31	c.1069A>T (p.Lys357Ter)	IDC	Negative	Negative	Negative	TNBC
IH4NMT	49	c.4834 C>T (p.Gln1612Ter)	IDC	Negative	Negative	Negative	TNBC
691NMT	56	c.2572C>T (p.Gln858Ter)	IDC	Negative	Negative	Negative	TNBC
IH11NMT	56	c.3607C>T (p.Arg1203Ter)	IDC	Positive	Positive	Negative	Luminal
1172NMT	73	c.5470_5477del (p.Ile1824Aspfs)	IPC	Positive	Positive	Negative	Luminal
1066NMT	34	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1213NMT	42	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1226NMT	44	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1000NMT	32	Non-carrier	DCIS	Positive	Positive	Negative	Luminal
1044NMT	46	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1048NMT	47	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1199NMT	53	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1112NMT	52	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1093NMT	59	Non-carrier	DCIS	Negative	Negative	Negative	TNBC
1485NMT	41	Non-carrier	IDC	Positive	Positive	Negative	Luminal
1068NMT	63	Non-carrier	DCIS	Negative	Negative	Negative	TNBC
1523NMT	76	Non-carrier	IDC	Positive	Positive	Negative	Luminal
1092NMT	44	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1283NMT	43	Non-carrier	IDC	Negative	Negative	Negative	TNBC

BRCA1 transcript: NM_007294.3

Abbreviations: NMT, normal mammary tissue. ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; IDC, invasive ductal carcinoma; IPC, invasive papillary carcinoma; DCIS, ductal carcinoma in situ.

Table S6. Clinico-pathologic information of the breast cancers from BRCA1/2 germline mutation carriers from The Cancer Genome Atlas (TCGA) database included in this study for comparisons.

Sample ID	Age	Gene	Mutation Type	Germline mutation	Histologic	ER statuses	PR statuses	HER2 statuses	PAM50 subtype
TCGA-AN-A0AL	41	BRCA1	Frameshift	c.5125_5126insC;p.Gln1709fs	IDC	Negative	Negative	Negative	Basal-like
TCGA-D8-A1XK	55	BRCA1	Frameshift	c.3390_3391insT;p.Phe1130fs	IDC	Negative	Negative	Negative	Basal-like
TCGA-E2-A14N	37	BRCA1	Frameshift	c.329_330insA;p.Lys110fs	IDC	Negative	Negative	Negative	Basal-like
TCGA-E9-A22E	56	BRCA1	Frameshift	c.5125_5126insC;p.Gln1709fs	IDC	Positive	Positive	NA	Luminal B
TCGA-E9-A22G	47	BRCA1	Frameshift	c.5125_5126insC;p.Gln1709fs	IDC	Negative	Negative	NA	Basal-like
TCGA-E9-A3QA	33	BRCA1	Frameshift	c.5125_5126insC;p.Gln1709fs	Medullary	NA	NA	NA	Basal-like
TCGA-A2-A0D2	45	BRCA1	Frameshift	c.4996delG;p.Val1666fs	IDC	Negative	Negative	Negative	Basal-like
TCGA-AN-A0FL	62	BRCA1	Frameshift	c.3894delA;p.Glu1298fs	IDC	Negative	Negative	Negative	Basal-like
TCGA-AO-A124	38	BRCA1	Frameshift	c.68_69delAG;p.Glu23fs	NA	NA	NA	NA	Basal-like
TCGA-AR-A24Q	49	BRCA1	Frameshift	c.782delG;p.Ser261fs	IDC	Positive	Negative	Negative	Basal-like
TCGA-E9-A244	54	BRCA1	Frameshift	c.3894delA;p.Glu1298fs	IDC	NA	NA	NA	Basal-like
TCGA-EW-A1P4	43	BRCA1	Frameshift	c.1740_1743delCAGT;p.Val580fs	Medullary	Negative	Negative	Negative	Basal-like
TCGA-BH-A0BL	35	BRCA1	Missense	c.181T>G;p.Cys61Gly	IDC	Negative	Negative	Negative	Basal-like
TCGA-C8-A12K	80	BRCA1	Missense	c.5095C>T;p.Arg1699Trp	IDC	NA	NA	NA	Basal-like
TCGA-D8-A147	45	BRCA1	Missense	c.181T>G;p.Cys61Gly	IDC	Negative	Negative	Negative	Basal-like
TCGA-A7-A4SE	54	BRCA1	Nonsense	c.913G>T;p.Glu305*	IDC	Negative	Negative	Negative	Basal-like
TCGA-AN-A0FX	52	BRCA1	Nonsense	c.3826C>T;p.Gln1276*	IDC	Negative	Negative	Negative	Basal-like
TCGA-E2-A14Z	64	BRCA1	Splice site	c.453A>C	IDC	Positive	Positive	Negative	Luminal B
TCGA-A2-A0T0	59	BRCA2	Frameshift	c.6942_6945delAATA;p.Thr2314fs	IDC	Negative	Negative	NA	Basal-like
TCGA-A8-A097	65	BRCA2	Frameshift	c.657_658delTG;p.Thr219fs	IDC	Positive	Positive	NA	Luminal B
TCGA-AO-A03V	41	BRCA2	Frameshift	c.5946delT;p.Ser1982fs	IDC	NA	NA	NA	Luminal A
TCGA-AR-A2LN	65	BRCA2	Frameshift	c.8533_8534delAG;p.Arg2845fs	IDC	Positive	Positive	Negative	Luminal A
TCGA-B6-A0I8	46	BRCA2	Frameshift	c.7683_7684delGT;p.Gln2561fs	IDC	NA	NA	NA	Luminal A
TCGA-BH-A1FN	34	BRCA2	Frameshift	c.2806_2809delAAAC;p.Lys936fs	IDC	Positive	Positive	NA	Luminal B
TCGA-C8-A8HR	49	BRCA2	Frameshift	c.8996_8997delTG;p.Leu2999fs	IDC	NA	NA	NA	Luminal B
TCGA-D8-A1JD	41	BRCA2	Frameshift	c.6847delC;p.Pro2283fs	IDC	Positive	Positive	Negative	Luminal B
TCGA-E2-A15S	34	BRCA2	Frameshift	c.1262delA;p.Gln421fs	IDC	Positive	Negative	Equivocal	Luminal B
TCGA-EW-A1PD	61	BRCA2	Frameshift	c.767_771delCAAAT;p.Thr256fs	IDC	Positive	Positive	Equivocal	Luminal B
TCGA-HN-A2NL	56	BRCA2	Frameshift	c.6642delT;p.Thr2214fs	IDC	Negative	Negative	Negative	Basal-like
TCGA-AR-A24L	26	BRCA2	Frameshift	c.1806_1807insA;p.Gly602fs	IDC	Positive	Positive	Equivocal	Luminal B
TCGA-BH-A0B4	65	BRCA2	Frameshift	c.8203_8204insC;p.Pro2735fs	IDC	Positive	Positive	Equivocal	Luminal B
TCGA-BH-A0BZ	59	BRCA2	Frameshift	c.9090_9091insA;p.Thr3030fs	IDC	Positive	Positive	Negative	Luminal B
TCGA-BH-A8FZ	58	BRCA2	Frameshift	c.1806_1807insA;p.Gly602fs	ILC	NA	NA	NA	Luminal A
TCGA-E9-A1NE	28	BRCA2	Frameshift	c.3975_3976insTGCT;p.Thr1325fs	IDC	Positive	Positive	Negative	Luminal B
TCGA-A2-A25E	34	BRCA2	Nonsense	c.9294C>G;p.Tyr3098*	IDC	Positive	Positive	Equivocal	Luminal B
TCGA-A8-A09A	40	BRCA2	Nonsense	c.5857G>T;p.Glu1953*	IDC	Positive	Positive	Negative	Luminal A
TCGA-A8-A09W	70	BRCA2	Nonsense	c.4552G>T;p.Glu1518*	IDC	NA	NA	NA	Luminal B
TCGA-BH-A0AZ	47	BRCA2	Nonsense	c.5864C>A;p.Ser1955*	IDC	Positive	Positive	Negative	Luminal A
TCGA-LL-A740	61	BRCA2	Splice site	c.517A>G	IDC	NA	NA	NA	HER2-enriched
TCGA-D8-A1Y0	65	BRCA2	Splice site	c.314_317delTAGG	IDC	Positive	Positive	Negative	Luminal B
TCGA-A2-A0SU	66	BRCA2	Nonsense	p.Y3308*	IDC	Positive	Positive	Negative	Luminal A

BRCA1 transcript: NM_007294.3

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; NA, not accessible.

Table S7. Clinicopathologic information of 7 BRCA1 mutation carriers and 10 non-carriers whose samples for immunohistochemistry staining.

Sample ID	Age	BRCA1 germline mutation	Histologic	ER statuses	PR statuses	HER2 statuses	Molecular subtype
ANM-1*	25	c.4013del (p.Lys1338Argfs)	IDC	Positive	Positive	Negative	Luminal
ANM-2*	43	c.2194G>T (p.Glu732Ter)	IMC	Positive	Negative	Negative	Luminal
ANM-3*	29	c.1069A>T (p.Lys357Ter)	IDC	Negative	Negative	Negative	TNBC
ANM-4*	43	c.5251C>T (p.Arg1751Ter)	IDC	Negative	Negative	Negative	TNBC
IH1NMT#	34	c.3288_3289del (p.Leu1098Serfs)	IDC	Positive	Positive	Negative	Luminal
1643NMT#	35	c.2899del (p.Thr967Leufs)	DCIS	Negative	Negative	Negative	TNBC
1550CNMT	46	c.1069A>T (p.Lys357Ter)	IDC	Negative	Negative	Negative	TNBC
1178NNT	49	Non-carrier	ILC	Positive	Positive	Negative	Luminal
1186NNT	50	Non-carrier	DCIS	Positive	Positive	Negative	Luminal
1197NNT	41	Non-carrier	IDC	Positive	Negative	Negative	Luminal
1211NNT	36	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1298NNT	37	Non-carrier	DCIS	Negative	Negative	Negative	TNBC
1353NNT	68	Non-carrier	IDC	Positive	Positive	Negative	Luminal
1141NNT	41	Non-carrier	IDC	Positive	Positive	Negative	Luminal
1099NNT	45	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1045NNT	62	Non-carrier	IDC	Negative	Negative	Negative	TNBC
1035NNT	53	Non-carrier	IDC	Negative	Negative	Negative	TNBC

BRCA1 transcript: NM_007294.3

* These samples were also been performed scRNA-seq.

These samples were also been performed bulk RNA-seq.

Abbreviations: NMT, normal mammary tissue; CNMT, contralateral normal mammary tissue ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; IDC, invasive ductal carcinoma; IMC, invasive micropapillary carcinoma; DCIS, ductal carcinoma in situ; ILC